DOCUMENT RESUME

ED 134 968 CS 003 232

AUTHOR	Fletcher, John D.; And Others
TITLE	Historical Antecedents and Contemporary Trends in Literacy and Readability Research in the Navy.
INSTITUTION	Navy Personnel Research and Development Center, San Diego, Calif.
SPONS AGENCY	Office of Naval Research, Arlington, Va.
REPORT NO	NPRDC-TR-77-15
PUB DATE	Jan 77
NOTE	107p.
EDRS PRICE	MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS	Adult Basic Education; *Adult Literacy; Composition Skills (Literary); Comprehension; Educational History; Information Processing; *Instructional Materials; *Literacy Education; Literature Reviews; *Readability; *Reading Research

ABSTRACT

These conference papers discuss the history of reading and readability research in the Navy, contemporary literacy research in the Navy, and research dealing with the readability and comprehensibility of written materials used in training or on the job. The first paper traces the interest in and development of literacy training in the Navy from its earliest days, focusing on the increasing importance of literacy in relation to technological advances. The second paper concentrates primarily on the spectrum of research being conducted at the present time, dealing with all facets of the literacy problem in the Navy. The third paper is an extensive survey of the literature and the state of the art in the assessment of readability and comprehensibility of written materials in the Navy. (Author/AA)

Documents acquired by ERIC include many informal unpublished materials not available from other sources. ERIC makes every effort to obtain the best copy available. Nevertheless, items of marginal reproducibility are often encountered and this affects the quality of the microfiche and hardcopy reproductions ERIC makes available via the ERIC Document Reproduction Service (EDRS). EDRS is not responsible for the quality of the original document. Reproductions supplied by EDRS are the best that can be made from the original.

ED134968

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

NPRDC TR 77-15

January 1977

HISTORICAL ANTECEDENTS AND CONTEMPORARY TRENDS IN LITERACY AND READABILITY RESEARCH IN THE NAVY

John D. Fletcher
Thomas M. Duffy
Thomas E. Curran

Reviewed by
John H. Steinemann

Approved by
James J. Regan
Technical Director

CS 003 232

Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
NPRDC TR 77-15

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
HISTORICAL ANTECEDENTS AND CONTEMPORARY TRENDS IN LITERACY AND READABILITY RESEARCH IN THE NAVY

5. TYPE OF REPORT & PERIOD COVERED
Final Report

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
John D. Fletcher
Thomas M. Duffy
Thomas E. Curran

8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Navy Personnel Research and Development Center
San Diego, California 92152

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
62757N
PF55.522.002

11. CONTROLLING OFFICE NAME AND ADDRESS
Personnel and Training Research
Office of Naval Research (Code N00014)
800 N. Quincy Street, Arlington, Virginia 22217

12. REPORT DATE
January 1977

13. NUMBER OF PAGES
100

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Comprehensibility
Education
Functional Literacy
Information Processing
Job Reading
Literacy
Instruction
Readability
Reading
Technical Documentation
Text Design
Useability
Writing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
This report is a compendium of three perspective papers presented at the Conference on Reading and Readability Research in the Armed Services held on October 28-30, 1975 in Monterey, California. The papers cover respectively, the history of reading and readability research in the Navy, contemporary literacy research in the Navy, and research dealing with the readability and comprehensibility of written materials used in training or on the job. The papers were considered to be of sufficiently wide interest in the Naval service that their publication as a technical report to be disseminated widely was

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

20. (Continued)

indicated. The Fletcher paper traces the interest in and development of literacy training in the Navy from its earliest days, focusing on the increasing importance of literacy with technological advances. The Duffy paper concentrates primarily on the spectrum of research being conducted at the present time, with emphasis on NAVPERSRANDCEN work, dealing with all facets of the literacy problem in the Navy. The Curran paper is an extensive survey of the literature and the state of the art in the assessment of the readability and comprehensibility of written materials in the Navy.

FOREWORD

The three papers that comprise this report were prepared under Exploratory Development Task Area PF55.522.002 (Methodology for Developing/Evaluating Navy Training Programs). The papers were presented by personnel from the Navy Personnel Research and Development Center at the Triservice Conference on Reading and Readability Research in the Armed Services held in October 1975. The Commanding Officer and Technical Director of NAVPERSRANDCEN were represented at that conference by Dr. Edwin G. Aiken.

Appreciation is expressed to Dr. Marshall J. Farr of the Office of Naval Research, who acted as technical monitor for the conference and its proceedings, and to the staff of the Western Division of the Human Resources Research Organization, who conducted the conference.

J. J. CLARKIN
Commanding Officer

SUMMARY

In October 1975, a triservice conference on Reading and Readability Research in the Armed Services was held at the Naval Postgraduate School in Monterey, California. The Navy was represented by the Navy Personnel Research and Development Center (NAVPERSRANDCEN) and the Naval Air Systems Command; the Air Force by the Human Resources Laboratory; and the Army by the Army Research Institute for the Behavioral and Social Sciences. The Personnel and Training Research Programs Division of the Office of Naval Research acted on behalf of the three services in contracting with the Human Resources Research Organization to conduct the conference.

Personnel from NAVPERSRANDCEN made three presentations at the conference. The first provided an historical perspective on literacy training in the Navy. The second presented an overview of current research on literacy in the Navy, with particular emphasis on recent work at the Center. The last addressed the area of readability and comprehensibility research in the Navy.

These papers are forwarded by this report since it is felt that the papers are of sufficiently wide interest in the Naval Service to justify extracting them from the overall Conference Proceedings for separate distribution.

CONTENTS

Page

INTRODUCTION 1

  Problem 1
  Background 1
  Purpose 1

HISTORICAL PERSPECTIVE ON LITERACY TRAINING IN THE NAVY 5

LITERACY TRAINING IN THE NAVY 29

READABILITY RESEARCH IN THE NAVY 69

DISTRIBUTION LIST 95

INTRODUCTION

Problem

The success of Navy education and training programs rests not only on assuring a systematic approach to instructional design and delivery, but also on improving the match between the entering skills of the students and the requirements of their curricula. Language skills, notably reading, are among the most fundamental prerequisite skills, covering virtually every course of Navy instruction and most Navy jobs. Closely related to reading skills is the readability of the materials dealt with in class and at work. Problems relating to both reading skills and readability have been documented many times over a wide span of years. There is a need to obtain both an historical perspective on Navy literacy problems and a contemporary look at Navy research and development in the area. The present report provides ample data and theory on all the principal issues.

Background

This report derives from a triservice conference held in October 1975 at Monterey, California. The three papers included in this report make up the contribution of NAVPERSRANDCEN personnel to that conference. Papers from the other services, plus remarks by consultants, can be found in the overall proceedings of the conference.[1]

The Center participants divided the area into an historical perspective on literacy training in the Navy, literacy research in the Navy, and readability research in the Navy. An effort was made to include consideration of long-standing and current problems, conceptual issues, and representative examples of data from ongoing Navy research programs.

Purpose

The purpose of this report is to make these papers available on a widespread basis. They are summarized below and presented in full in the following sections.

"Historical Perspective on Literacy Training in the Navy"

The first paper was prepared by Dr. John D. Fletcher. The author divides Navy history on this subject into two areas; before and after 1943. Up to 1943, the increasing importance of literacy among enlisted personnel is related to technological advances, the duties of the Chaplain Corps, the appearance of enlisted handbooks, the use of written screening tests, and literacy training in a Navy prison rehabilitation program. After 1943, special in-service programs began to appear. These became progressively more sophisticated in terms of curricular content, individualization, student selection, evaluation, and supporting research. Fletcher makes it clear that support for such programs varied directly with quality declines in the manpower pool, response to special programs such as Project 100,000 and, most particularly, mobilization. The author feels that continued support of literacy training by the Navy depends on the development of the concept of job-related or functional literacy training.

---

[1] Sticht, T. G., and Zapf, D. W. (Eds.). Reading and Readability Research in the Armed Services.

"Literacy Research in the Navy"

The second paper was prepared by Dr. Thomas M. Duffy. The author begins by characterizing current concerns with literacy levels in the Navy as turning on the apparent large discrepancy between personnel reading ability and the reading difficulty of much Navy technical writing. He then proceeds to present data from two recent investigations carried out at the Navy Personnel Research and Development Center. The first, a study that compares reading skills with the difficulty of recruit and apprentice school reading materials, shows that there is a clear skill deficit in about 1 in 5 recruits. Further analyses relate recruit reading levels to basic test battery scores, amount of education, race, and the probability of attrition from recruit training. The second study examines reading skills in a sample of Navy Class "A" School students. When reading skill was compared to the reading difficulty of the manuals written for each of the "A" School ratings, some schools were found to have large numbers of men with reading deficits. Further analyses indicated that weekly test scores in some "A" Schools are more accurately predicted from a knowledge of a trainee's reading ability than from his basic test battery scores or a nonverbal measure of general intellectual ability. Reading skill appears to be most predictable in schools with a high proportion of low-ability readers. Duffy continues his paper with a review of current Navy/Marine Corps literacy training programs. Four programs are compared on several student and instructional variables, and it is shown that, despite wide differences on these variables, the reading gain scores for all programs are approximately the same. The reasons for this and means for increasing the impact of literacy training programs in the Navy are then discussed. The paper concludes with some preliminary data on predictors of success in reading training.

"Readability Research in the Navy"

The last paper was prepared by Dr. Thomas E. Curran. The author reports on his state-of-the-art survey on the readability and comprehensibility of technical materials. First, he compares the distribution of Navy recruit reading skills with the distribution of the reading difficulty of 185 Navy rate training manuals. A clear mismatch is apparent, with difficulty exceeding ability. Next, he considers the problems associated with measuring the readability of text and describes the major formulas and procedures. Particular attention is given to the problem of technical terminology in such measures and to the distinction that must be drawn between readability and comprehensibility. Curran points out that, although the readability of a passage can be adequately indexed from existing formulas dealing with "countable" features of the writing, as of this time, only the performance of a group of subjects on a post-reading test can determine its comprehensibility. He concludes this section with a discussion of means for specifying the reading requirements of Navy jobs. Curran next reviews research on the production of readable writing. Advice from writers' style guides is compared and found to be in considerable disagreement. Curran presents a diagram of the iterative relationship that should hold between the prediction and the production of readable writing. The paper concludes with a review of existing and developing procedures for automating the prediction process. These include the Navy Automated Counter; the Automated Readability Index; the Reading Ease Assessment Device; the Computer Aided Revising, Editing, and Translating System; and the Technical Review and Update of Manuals and Publications System.

HISTORICAL PERSPECTIVE ON LITERACY TRAINING IN THE NAVY

J. D. Fletcher

History seems to be not so much a record of what has happened as a record of what we remember, and any historical perspective is most probably a perspective on perspectives. This is certainly true of the current historical perspective on literacy training in the Navy. Because of the subjective nature of the data that support it, this perspective is divided into two parts: a very brief part dealing with all of Navy history prior to 1943 and a more extensive part beginning in 1943 and ending with the present. Further, very little information is presented on readability and technical writing. Historical information on these topics doubtless exists, but it is buried in the instructions and official correspondence of the Navy Department.

A comprehensive, thoroughly satisfactory definition of literacy training was not attempted in this report. A rough description of literacy training as an attempt to bring personnel who read below the 4.0 grade level up to that level appears to serve fairly well the needs and scope of this report. In the history of Navy training, this description emphasizes training for enlisted men who are in their first tour of duty.

Literacy Training Before 1943

Enlisted personnel are, at best, neglected in military histories. Battles, tactics, and technologies tend to be recorded in some detail, but the ability and character of the enlisted force that are basic to any military service are often obscured by the statistics of manpower supply, loss, and demand. If it is reasonable to assume that enlisted personnel acquire a greater significance as the technological demands of their duties increase, then their neglect in histories of the United States Navy is particularly regrettable. It is difficult to read any account of enlisted life without bringing away an impression that Navy duties have always demanded substantial technological capabilities. The introduction of steam and electrical systems aboard ships obviously increased the technological demands of enlisted Navy jobs as Cummings (1929), Harrod (1973), Potter (1913), and others have indicated, but discussions, such as those of Luce (1890) and Niblack (1891), of training problems that existed in the earlier, sailing fleet indicate that substantially more than disciplined responses to orders and knowledge of nautical terminology were relevant objectives in transforming "landsmen" into seamen. Most of this training was accomplished on-the-job following the apprentice, journeyman, master craftsman model. It required listening but not reading skills, and it continued throughout a sailor's career. Despite the predominance of this mode of training, there were significant efforts as early as the first half of the 19th century to provide instruction in the basic skills of reading, writing, and arithmetic to Navy personnel. There was no land-based training establishment at this time and basic skills training took place entirely aboard ship, at dockside and while the ship was underway. There appear to be two primary reasons for the early literacy training: the religious reform movements of the 1700's and the assignment of teenage boys to the fleet.

Religious Reform

In the late 17th century and during much of the 18th century, a major transformation in Western civilization took place under the auspices of the Enlightenment. Leaders of this movement placed great faith in human reason backed by the findings of science and submitted to question all authority and absolute standards. As a result, a sympathetic and humanitarian outlook on the condition of all men was stimulated, especially by religious manifestations of this movement. Protestantism particularly fostered an enlightened self-interest on the part of its followers by emphasizing good works and the notion that every man is his brother's keeper. Heaven would be won and past wrongs atoned for, if each man would firmly embrace religion and help his fellow man, especially his less fortunate fellow man, to do the same. Some members of the Evangelical movement, which grew naturally from the Enlightenment, chose to focus their interest on the sailors of the naval and merchant service. Revivals were held, Sunday schools were conducted, and tracts were distributed. With the distribution of the tracts came the realization that the Word of God was inaccessible to many sailors because they could not read. Moreover, sailors' moral welfare appeared endangered because their lack of arithmetic skills made them easy prey for the peddlers of prurience who were naturally drawn to the full pocketbooks and poorly educated minds of newly disembarked sailors. So it was that the first to minister to the need for essential skills training among Navy enlisted men were the early Navy chaplains (Langley, 1967).

The regulation of 1802 described the chaplain's duties as the following:

> 1. He is to read prayers at stated periods; perform all funeral ceremonies over such persons as may die in the service, in the vessel to which he belongs; or, if directed by the commanding officer, over any person that may die in any other public vessel.
>
> 2. He shall perform the duty of a schoolmaster; and to that end he shall instruct the midshipmen and volunteers, in writing, arithmetic, and navigation, and in whatsoever may contribute to render them proficients. He is likewise to teach the other youths of the ship, according to such orders as he shall receive from the captain. He is to be diligent in his office and such as are idle must be represented to the captain, who shall take notice thereof. (Burr, 1939, p. 111).

In addition to chaplains, there were schoolmasters and teachers in the Navy (as evidenced by the law of March 1799 assigning them three-twentieths of prize money, which was approximately the amount assigned to warrant officers) but, in general, the schoolmasters made a poor showing. Allegations of sloth and drunkeness on their part appear well founded. Schoolmasters' duties varied with their commanders, but they were generally charged with instructing the boys, apprentices, and midshipmen assigned to ships. The employment of schoolmasters was never extensive in the Navy, and it gradually died out. However, the history of Navy chaplains is far more honorable than that of the schoolmasters, and the Chaplain Corps maintained its interest in essential skills instruction for enlisted men throughout the history of the Navy (Drury, 1949). There was, of course, little systematic method in this instruction, and its success most probably keyed on the moral rather than cognitive well-being of its students.

Apprentices

From the beginning of the Navy, there were boys assigned to ships. Attempts to standardize their employment and treatment were embodied in the apprentice systems set up in 1837, 1855, 1863, and 1875. Of all the training ventures before the Spanish-American War, only the naval apprentice program provided a prototype for the modern Navy (Harrod, 1973). Despite this fact, the apprentice system suffered unpromising beginnings. The 1837 effort failed because apprentices' hopes of obtaining commissions were not fulfilled and because of an alleged mutiny in 1842 on the training ship SOMERS in which a son of the Secretary of War was hanged, probably wrongfully (Langley, 1967). The 1855 program was interrupted by the outbreak of the Civil War. The 1863 effort was unsuccessful because apprentices were again disappointed in their hopes of obtaining commissions. However, the apprentice system continued to be revived because of very serious problems the Navy was experiencing in manning American ships with native-born seamen (e.g., Luce, 1874). In 1875 the Navy began enlisting boys 14-18 years of age to serve as apprentices until age 21. The Secretary of the Navy emphasized that it was not the object of the system to prepare boys for commissioning; the apprentices were to be trained in all the duties of sailors on a man-of-war, and, significantly, they were to receive an elementary English education (Harrod, 1973). Despite numerous problems, including the neglect of academic training, the system, with its commitment to literacy training, survived until it was replaced in 1904 by an officially established apprentice seaman rating with a minumum age requirement of seventeen. With this system evolved the prototype for today's recruit training.

Recruit training was essentially motivated by a need for standardization in training landsmen and apprentices for shipboard duty. A natural means for standardizing training that was employed very quickly after the need for it was recognized was the publication of drill books. Fullam's Recruit's Handy Book and McLean's Bluejacket's Manual both appeared in 1902; Fullam's book lasted until the 1920's and McLeans' manual is currently in its nineteenth edition. By General Order 114 of November 1902, the Navy Department required all recruits to know the contents of the Recruit's Handy Book and issued a copy to each. Fullam also published in 1902 the Petty Officer's Drill Book.

The appearance of these handbooks signaled the fact that oral instruction was no longer sufficient for an enlarged Navy (cf. Harrod, 1973). Presumably this fact was already apparent. Some of the technical equipment aboard Navy ships at the beginning of the century was fairly sophisticated, and it seems likely that this equipment was accompanied by essential manufacturer's literature on its operation, maintenance, and repair. However, there appears to be little record of the nature, preparation, and supply of this technical documentation. In any case, the appearance of the drill books signified an official expectation that seamen would be able to read, and literacy effectively became a requisite Navy skill.

Despite a national requirement for literacy, about 11.3 percent of the United States population in 1900 could be classified as illiterate (Harman, 1970). There were no systematic attempts to screen Navy recruits until 1925 (Harrod, 1973), and illiteracy among Navy personnel became increasingly serious. Although there is no direct reference to this effect, the systematic screening that began in 1925 included written tests, successful performance on which required literacy. It seems likely then, that the incidence of illiteracy in the Navy was substantially reduced by the use of these tests.

Aside from the work of the chaplains, the only formal literacy training that appears to have been supported by the Navy in the period 1900-1943 was in Portsmouth Prison where some effort was made to rehabilitate men who were classified and jailed as deserters because of their inability to read furlough orders (Potter, 1918). These men would go on leave and return on what they thought was the appropriate date only to find themselves scheduled for court-martial.

Despite the screening process begun in 1925 and official policies for their exclusion, illiterates continued to appear in Navy billets throughout the entire period (Special Training Program, 1951). Other than the rehabilitation program at Portsmouth Prison, no official efforts were made to help these men achieve literacy. There was, and is, considerable unskilled labor required in many fleet billets, and these men were typically assigned to do this labor. As might be expected, rates of promotion and reenlistment among these men were not high. During World War II, manpower supply began to run substantially short of demand and in June 1943 the Navy reluctantly agreed to accept its fair share of illiterates under Selective Service.

Literacy Training in 1943 and After

The Special Training Program

Although, prior to June 1943, there was no official recognition of the fact that room would have to be made for enlisted personnel who could neither read nor write, a great many men in this category were already serving in the Navy. Immediately after Pearl Harbor, recruiting stations were released from the obligation to administer the General Classification Test (GCT) to applicants and, since this regulation remained in abeyance until the start of 1945, many illiterates found their way into the Navy through traditional recruiting channels (Special Training Program, 1951). In August 1945, the Director of Training indicated some of the problems the Navy and illiterates experienced with each other in the following summary:

(1) At times the period allowed for recruit training was contracted by the demands of the service to four or five weeks. Under these circumstances the trainee was obliged to acquire a large part of his instruction through reading. It was found that it took approximately four times as long to train an illiterate to perform an average Navy job as it did one who could read.

(2) The establishment of a training program which did not depend on the use of printed matter would have been both difficult and expensive. Experience showed that it was simpler and more economical to teach men to read than to devise materials which did not require this knowledge.

(3) The establishment of a smooth administrative routine was grossly complicated by the presence of non-readers. A system for the rapid handling of records was a virtual impossibility where men could not fill out information blanks, pay recipts, proficiency slips, allotment cards, etcetera.

(4) Sufficient education to read safety precautions was essential for men working with machinery, high explosives, and heavy cargo. Serious accidents were traced directly to men's inability to read warnings and study safety instruction.

(5) A social barrier of serious implications was found to exist between literate and illiterate personnel.

(6) The administrative dualism resulting from putting literates and illiterates together caused confusion. Literates tended to resent the long oral directions which they had to listen to for the sake of the illerates in their number.

(7) A very large number of minor disciplinary problems were the direct outgrowth of misunderstandings caused by inability to read station orders, watch bills, leave and liberty regulations, and safety instructions.

(8) An inability to read and write letters constituted among illiterates a serious morale problem and consequent obstacle to satisfactory adjustment to naval life. It became increasingly evident that a knowledge of reading and writing helped to overcome a feeling of inferiority and tended to develop initiative, aggressiveness, and more willing acceptance to the conditions of military life. (Special Training Program, 1951, pp. 2-3)

Despite this analysis which was written after the fact, there is substantial evidence that authorities were reluctant to accept responsibility for setting up a special literacy program long after the need for such a program became obvious. However, as the situation gradually worsened, the Navy stumbled into a situation that required recognition of the need for a special literacy program. On 30 September 1943, the Naval Training Section in Norfolk, Virginia reported that illiterates were being received from boot camps in increasing numbers. Since all the facilities at this section were need for the destroyer and destroyer escort training programs, permission was requested to transfer these personnel to the Naval Training Section at Bainbridge, Maryland for further instruction. Evidently, the permission was granted without due consideration of the practical consequences of this decision. The result was that Bainbridge found itself deluged with illiterates and appealed to the Bureau of Naval Personnel for help. The immediate response of the Bureau was to cancel the permission that had previously been granted to Norfolk. However, as a result of these events, the need for special literacy training was brought to command attention (Special Training Program, 1951).

The presence of a growing body of illiterates in the Navy created a training problem that it was totally unprepared to face. Luckily, as a thorough review of World War II literacy training programs by Fattu, Mech, and Standlee (1953) shows, the foundation for solving this problem had been laid much earlier by the work of the Civilian Conservation Corps (CCC).

The CCC was established in March 1933 primarily as a means for providing productive employment for young men during the depression. Although it was administered by the Army throughout its 9-year existence, much use was made of professional educators in developing and guiding the educational phases of the program. The CCC education program was secondary to the work program, and the literacy program was only a small part of the education program. It was created to serve the estimated 6 percent of CCC enrollees who were functionally illiterate in the camp life situation (Couch, 1944). Nevertheless, the literacy training program did exist as an identifiable entity, and it served as a foundation for the development of similar programs, first by the Army in 1941 and later by the Navy in 1944 (Fattu et al., 1953).

The Army program was discussed by Heath (1946), who pointed out that there were three distinct groups of illiterates targeted by the Army program: English-speaking illiterates, non-English-speaking illiterates, and Oriental literates. Training was tailored to the special needs of each of these groups. The teaching program for the English-speaking illiterates passed through five distinct phases. These phases were sequenced and a student had to master each phase before proceeding to more advanced ones. Heath describes the phases as the following:

1. Consonants and Key Words. Sounds of the consonants were taught in this phase by associating 21 key words with English consonants.

2. Monosyllables and Vowels. Nonsense syllables such as ter, ker, nub, etc. were initially taught in this phase in an effort to train students in the relationship between orthography and sound.

3. Introduction of Polysyllables. Students were taught to analyze polysyllables into monosyllables.

4. Sentence Structure and Word Recognition from Context. Further instruction in word attack was presented and instruction in constructing sentences and learning (recognizing) words from context was added in this phase.

5. Composition and Expression. Students were taught to read and prepare military messages and personal letters in this phase. They were also taught the use of such basic resources as dictionaries and telephone directories.

Discussion of the Army's literacy program is relevant because, when the Navy finally faced the need to produce a literacy training program, it turned in September 1943 to the Army for guidance, and the workbook material developed for the Navy was based on the five phases discussed above.

On 22 December 1943, the imminent appearance of a special literacy program for recruits was announced by the Navy. On 23 December 1943, the Naval Training Section at Great Lakes, Illinois was directed to prepare for the arrival on 3 January 1944 of an initial draft of 420 white illiterates. In March 1944, two Navy programs for literacy were established: one at Camp Peary, Virginia for Whites and one at Great Lakes for Blacks. Plans for the Camp Peary program called for a weekly input of up to 500 trainees with a total capacity of 6000. In fact, the total enrollment at Camp Peary quickly grew to 10,000 in April 1944 (Special Training Program, 1951). Notably, both these programs were set up as an integral part of recruit training; no programs were established for illiterates who were already in the Navy. At no time did the Bureau of Personnel formally accept responsibility for training illiterates who were above the recruit level; however, informal support was given to commanders who wished to aid illiterates under their command with the dissemination of literacy training materials throughout the fleet.

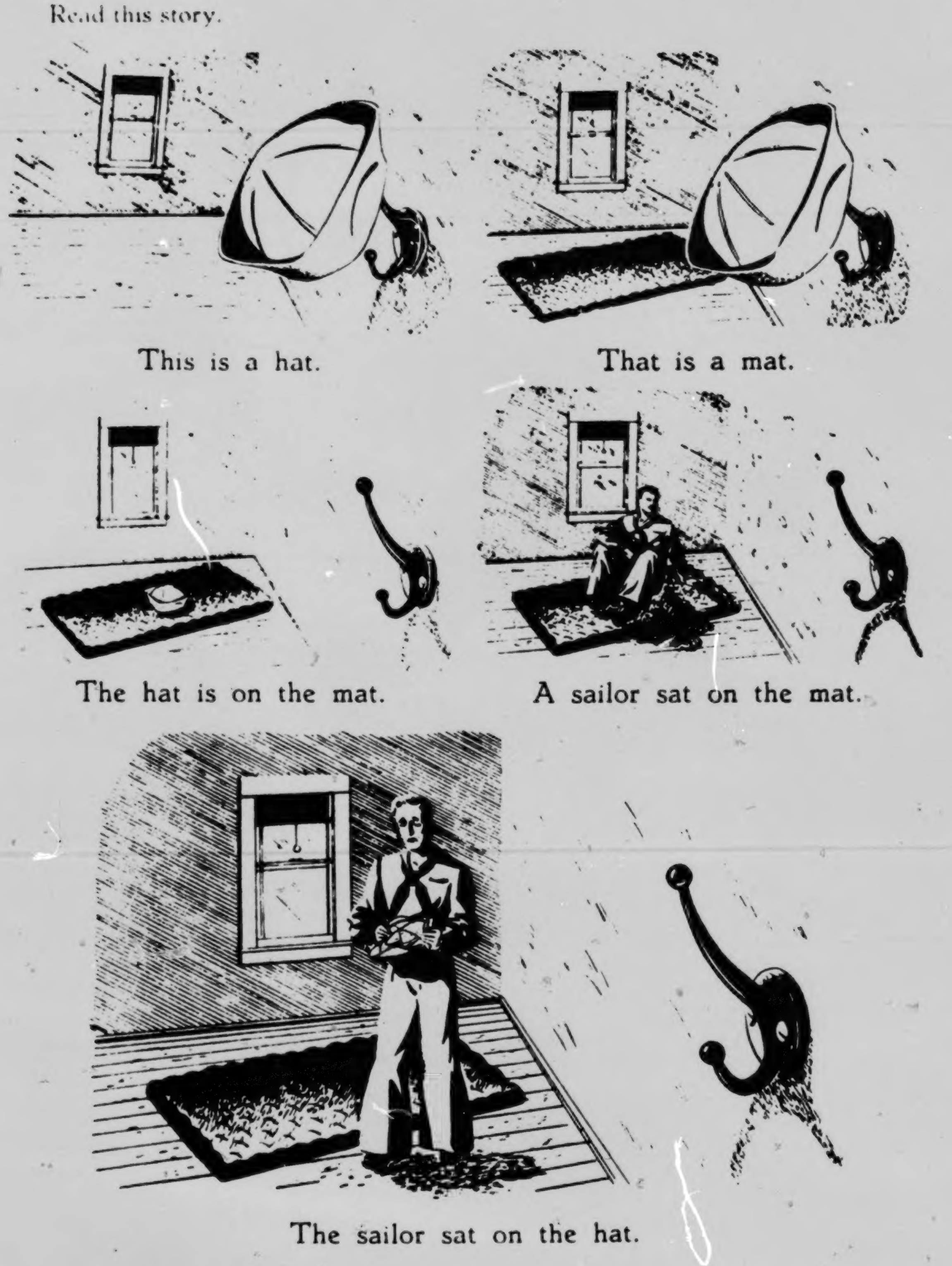
The curriculum that was initially devised for the literacy programs in early 1944 provided for only 133 hours of instruction in reading and writing and 73 hours of instruction in arithmetic, for a total of 205 hours of instruction out of the 576 hours originally called for. However, by January 1945 a considerably expanded and improved program had evolved. Four basic assumptions not previously annunciated determined the form of the Special Training Program, as it came to be called. First, it was a training program for adults. Although unable to read and write, the trainees came to the Navy having command of a well-established oral vocabulary together with a fund of experience that put them beyond the appeal of grade school readers. Second, the trainees were the products of a wide variety of environments, so that the only interests they could be expected to have in common would grow out of their life in the service. Third, the limitations on time allowed for acquiring literacy made it mandatory that the program be rigidly functional in nature. A nominal proficiency grade of 5.0 in reading and writing was established as the teaching goal, but the purpose behind the program was simply to qualify men to read, write, and figure sufficiently well to perform all essential Navy duties. Thus, it was expected that graduates of the course would be able to read watch lists and safety precautions, and would be able to fill in beneficiary slips and small stores chits. On the other hand, there would be no attempt to indoctrinate trainees formally in such refinements as capitalization, spelling, and punctuation. Fourth, it was assumed that the vast majority of instructors would be inexperienced in the type of teaching expected of them. They could be expected to rely unduly on lecture and blackboard methods and to encourage parrot-like memorization of lists of words without developing the skills of analysis and synthesis that are essential to literacy (Special Training Program, 1951). The CCC literacy materials were called the Camp Life series and the Army materials were called the Army Life series, so naturally the workbooks, basic readers, supplementary readers, tests, and teacher's manual developed for the Navy were called the Navy Life series. Private Pete was replaced by Seaman Sam.

Development of the program materials was described by Ross (1946):

> In writing this program, some radical departures from the conventional were taken, because the situation and the nature of the students and instructors demanded them. For example, a "reader" in the hands of an untrained instructor at the outset of the program would result in a static classroom situation in which the students "read" orally in rotation, with prompting, until the page has been virtually memorized. In the Navy Life series, therefore, the first book is not a "reader," but a workbook-type text which forces the instructor down from the platform among the students. The first of the readers is not introduced until considerable reading ability is developed through chart, blackboard, and workbook reading experiences. When it is introduced, no new skills or words are required for some time, and the student can read it easily for meaning. As a result, the student is literate so far as the readers are concerned from his very first experience with them.
>
> Comic books are exceeded in popularity by no other reading material. So, later in the program, when they can be handled easily, comic books are introduced for rapid, supervised classroom suplementary reading. These are regular commercial comic books, carefully selected in advance of publication, and then rewritten with a core vocabulary basic to the Navy Life program. (p. 204)

Notably the cornerstone of this program was not the basal readers, which was probably the case in civilian initial reading instruction, but the program's two workbooks functionally entitled Navy Life Book I and Navy Life Book II. Both books were constructed on a basic vocabulary that derived first from words shown by research (what specific research is not in the record) to be an essential minimum for literacy, and then, to as great an extent as possible, from naval terminology. Fans of Leonard Bloomfield will be interested in Figure 1, which is the first "story" that occurs in Navy Life Book I. The emphasis on the single spelling pattern "at" embedded in hat, mat, and sat will seem very familiar to those acclimated to Bloomfield's "Nan can fan Dan" (cf. Bloomfield & Barnhart, 1961). Extensive reliance was placed on illustrations to indicate word meaning, but other devices such as a visual acuity test and illustrations of phonetic similarity were also used. It was assumed that, by teaching reading and writing simultaneously, growing proficiency in either would increase proficiency in the other. The material in Navy Life Book I, which contained 400 illustrations, took as its common denominator barracks life, which all trainees might be expected to have in common. In the second volume, the emphasis was gradually shifted from phonetic elements to syllables and from illustrations to context as a means of furnishing clues to word meaning. The rigid control over the material that characterized Navy Life Book I was gradually relaxed in Book II, and was finally loosened so that any student capable of handling it would be able to conduct independent reading outside the classroom. Navy Life Book II was also prepared with the intention that it should be a useful adjunct to regular recruit training, since it contained a good deal of incidental information on such subjects as semaphore, firefighting, elementary navigation, naval customs, guns, ship types, seamanship, naval terminology, and personal hygiene.

As it evolved, the curriculum was developed on a flexible scale operating between the limits of 12 and 20 weeks. In other words, trainees received periodic tests starting with the eleventh week, and could be graduated any time thereafter by showing that they had achieved the required level of literacy. The minimum overall time prescribed for the course was 256 hours, and the maximum time was 528. Within this framework the variation in the allowance for reading and writing ran from 129 to 312 hours, and for arithmetic, from 63 to 118 hours.



Read this story.

This is a hat.

That is a mat.

The hat is on the mat.

A sailor sat on the mat.

The sailor sat on the hat.

Figure 1. The first story in the Navy Life workbook series.

Selection for the program depended on scores achieved on three tests:

1. The General Classification Test (GCT)--a test of general ability which in 1945 was reintroduced and given to all incoming personnel.

2. The Literacy Test (LIT)--a direct measure of reading achievement consisting of 44 items divided into three subtests: word recognition, sentence reading, and paragraph interpretation.

3. The Nonverbal Classification Test (NVCT)--a reflection of the GCT administerd in nonverbal terms through the use of pictures and geometric forms. The NVCT consisted of 75 items divided into five subtests: picture analogies, figure analogies, picture classification (opposites), figure classification (opposites), and matrics (comparable to Raven's matrics).

According to Hodges (1964), the following screening procedure was used:

1. All recruits scoring less than a standard score of 35 on the Reading Test (the GCT, when the Reading Test was discarded) of the Basic Test Battery were administered the LIT.

2. Men scoring 40 (equivalent to 5.0 grade level) or better on the LIT were considered "literate" and sent on to regular recruit training, and those scoring below 40 on the LIT were given the NVCT.

3. Men scoring higher than 34 on the NVCT were considered trainable and sent to the Special Training Program. (A raw score of 34 on the NVCT is between 30 and 35 Standard Score on the GCT.)

4. Men scoring below 34 on the NVCT were sent to a psychiatric unit for closer examination; they were suspected of being both illiterate and untrainable. (N. B. - The LIT and NVCT scores of 40 and 34 were later changed to 37 and 38 respectively as a result of an unidentified study, all records of which have vanished.)

In other words, the men sent to the special training units were those in need of literacy training (low LIT score) and judged capable of assimilating it (high NVCT score).

The Navy Reading Achievement Examination (NRAE) was also developed as a standard measure to determine when students in the variable-length Special Training Program were ready to graduate. Both the LIT and the NRAE were calibrated against the Gates Reading Survey to establish grade levels. The NRAE was designed to be functional and tested students for understanding Navy situations in which they might find themselves.

There appears to be a single evaluation study of the Navy's World War II Special Training Program. This study was administered and documented by Hagen and Thorndike (1953) on the basis of personnel records salvaged from the Naval Records Management Center, Garden City, New York. Two types of data were abstracted: (1) background facts that might predict success in the Navy and (2) facts about the man's career that might indicate his success in Navy duty assignments. Three groups of men were identified:

1. The "Illiterate" Group. This group consisted of 1026 men inducted into the Navy during August and September 1944 and initially assigned to the Special Training Program at Camp Peary.

2. The Control Group. This group was selected by taking a serial number that was five digits above the serial number of a member of the illiterate group. The man thus identified was included in the control group if he was not black, not illiterate, and came from the same geographical area as his illiterate counterpart. If a serial number five digits higher did not identify a qualified counterpart, one five digits below was selected and so on. In this way, a control group of 1021 men was chosen. This group resembled the illiterate group on most geographical and demographic measures.

3. The Marginal Group. This group was composed of 999 men who had scored below 36 on the GCT and who entered the Navy at about the same time as the illiterates. This group was geographically and demographically dissimilar from both the illiterate and control groups.

Hagen and Thorndike described the average member of their illiterate group as follows:

> He was inducted in the Navy when he was about 19 years old. At the time of induction, he was single but had one or more people who were partially or wholly dependent on him for support. Before induction, he had lived in a rural area of the South where the standard of living was below the average for the nation as a whole. He had completed the fourth grade in school and left school at the age of fourteen after having repeated at least three grades. Since leaving school, he had worked for his parents or a relative on a non-mechanized farm. In his spare time he hunted or fished. He made a score of 4 out of a possible 17 on the Qualifications Test and a score of 31 on the Navy General Classification Test. (p. 18)

Hagen and Thorndike summarized the differences they found between the illiterate and control groups as follows:

1. The illiterates were much more likely than were the control cases to be assigned to construction battalions, and were less likely to be assigned to U. S. permanent party or to auxiliary ships.

2. The illiterates tended to receive a lower average proficiency in rate. Only 50 percent received an average of 3.5 or over, as compared with 73 percent in the control group.

3. The illiterates received fewer promotions. Only 15 percent made petty officer, as compared with 37 percent in the control group.

4. The illiterates received more disciplinary actions. In the illiterate group, 23 percent had records of some type of disciplinary action, as compared with 11 percent of the control group. General courts martial were 10 times as frequent in the illiterate group.

5. The illiterates more frequently lost time due to misconduct—20 percent versus 7 percent.

6. The illiterates less frequently received an honorable discharge—83 percent versus 88 percent.

7. The illiterates were somewhat more likely to receive a medication survey, and were the only group to be surveyed for inability to learn--19 percent versus 15 percent.

8. The illiterates were somewhat more likely to incur a venereal infection—5 percent versus 3 percent.

9. The illiterates were slightly more likely to generate a Veterans Administration disability claim—11 percent versus 9 percent.

In most of the factors that distinguished the illiterate from the control group, the marginal group occupies an intermediate position, usually nearer the illiterate group than the control group. The only exception to this is the case of disciplinary actions; the marginal group were more often in trouble and their offenses were more serious.

The Hagen and Thorndike study is interesting, but it is essentially a study of personnel at different mental levels. It does not have much to say about the effectiveness of literacy training. It does, however, tell something about the success of illiterates in the Navy. In general, it seems reasonable to conclude, with the authors, that many or most of the illiterates appeared to make an accpetable adjustment to the Navy.

After the war, screening of illiterates with the GCT, LIT, and NVCT continued, but the Special Training Program was discontinued in the course of general demobilization. Ginzberg and Bray (1953) estimated that 35,000 men were assigned to the Navy's Special Training Program in the course of its history.

Recruit Preparatory Training

In 1950, with the influx of volunteers during the Korean War, commands were once more authorized and encouraged to identify and train any illiterates who might be aboard. By the spring of 1951, literacy training was being conducted informally in after-hours programs at the three recruit training centers (Bainbridge, Maryland; Great Lakes, Illinois; and San Diego, California). By September 1951, literacy training was planned on a full-time basis under the new title Recruit Preparatory Training (RPT) (Standlee, 1954). Officially, the main objective of the program was "to teach selected recruits to read and understand instructions and to prepare them to absorb military training. Recruits should attain a level of proficiency in reading which is comparable to completion of the fourth grade" (Curriculum for Recruit Preparatory Training, 1953, p. 5). Further, an accepting classroom atmosphere was to be created, as indicated by the following guidance provided by the Bureau of Personnel:

> 1. Make the first reading tasks simple enough to insure that the recruit experiences an early feeling of success in learning to read.
>
> 2. Show the recruit that his low reading achievement is a handicap which must be overcome in reaching his goal--becoming a useful sailor.
>
> 3. Create a permissive classroom atmosphere and start where the learner is. The individual recruit is the one who has to learn. The most the instructor can do is guide and help him. The instructor should guard against a highly directive, subject-matter-centered approach which is on a level that never reaches some recruits.
>
> 4. Integration of military subjects with reading, writing, and arithmetic. Some examples of how this can be done are listed below:
>
> > a. The instructor includes some topics related to service in the Navy at appropriate times in his reading, writing, and arithmetic classes. This procedure will probably be more effective during the latter part of RPT.
> >
> > b. Use of visits to ships or parts of the training center. Trainees can can write about what they see, read accounts to each other, learn Naval terms, etc.
> >
> > c. Learn to read safety signs found on the center or aboard ship.
> >
> > d. Preparation of RPT news sheet and use of local newspapers and announcements of recreation.
> >
> > e. Provide guidance in writing letters home about barracks, field-day, bunks, meals, training activities, etc.

f. Familiarization with types of material found in The Bluejackets' Manual, libraries on the center and aboard ship, and courses available through USAFI. (Curriculum for Recruit Preparatory Training, 1953, p. 11)

The curriculum materials were, with minor updating, composed of the Navy Life series prepared for the World War II Special Training Program, although use of other supplementary materials was encouraged (Standlee, Fattu, & Auble, 1954a). The selection procedure for RPT was the following, as described by Hodges (1964):

1. All men scoring below 36 on the GCT were administered both the NVCT and LIT.

2. Men scoring 38 and above on both tests were considered literate and trainable and sent to regular recruit training.

3. Men scoring 37 and below on the LIT (38 on the LIT is about a 4.7 grade level) and scoring 38 or higher on the NVCT were considered illiterate but trainable and sent to special training.

4. Men scoring less than 38 on the NVCT and on the LIT were suspected to be untrainable and were sent to a neuropsychiatric unit for further examination. From the neuropsychiatric unit, men were sent to special training, regular recruit training (on a trial basis), or discharged.

Referrals could also be made to the Special Training Program by Company Commanders in regular recruit training. The NRAE with a new form added was restored as an end-of-course criterion. The 5.0 grade level was still the course objective. Students were given 7-9 weeks to complete the course. An observer who was sent to the three RPT centers in 1953 reported that, in general, the selection and curriculum procedures in use followed the Bureau's directives, that a total of 239 trainees were enrolled in all three centers, that 85 to 94 percent of the trainees successfully completed the RPT program, and that they took 6-7 to 9 weeks to complete it (Fattu, Fay, D'Amico, & Standlee, 1954).

Two evaluative studies of RPT were completed. A study by Cofer (1954) investigated the effectiveness of RPT in achieving its most immediate objective, preparing the trainee for regular recruit training. Standlee (1954) dealt with the more ultimate criterion--performance of RPT graduates in fleet assignments.

A principal concept in Cofer's study is that of adjustment, specifically adjustment of RPT graduates to regular recruit training. For the purposes of this study, mental health considerations were set aside in favor of adjustment that emphasized, first, adequacy of performance of recruit duties, and, second, attitudinal and motivational factors in relationship to the Navy, recruit training, and civilian plans. Fourteen rating scales were developed to assess adjustment in the first sense, and an 80-item questionnaire was developed to assess adjustment in the second sense. Two groups of subjects were identified, one at an early stage of recruit training (Group I) and one at a later stage of recruit training (Group II). Within each of these groups, four subgroups were identified:

1. RPT graduates: Group I: N = 30
Group II: N = 21

2. Recruits with GCT scores below 35, but not RPT graduates:

Group I: N = 13
Group II: N = 18

3. Recruits with GCT scores between 35 and 40 inclusive:

Group I: N = 29
Group II: N = 46

4. Recruits with GCT scores above 40:

Group I: N = 30
Group II: N = 49

Each subject was evaluated on his adjustment to regular recruit training by peers and Company Commanders separately on the 14 scales reflecting adequacy of performance in recruit training.

In general, members of the RPT subgroup resembled the members of the high GCT subgroup on the criterion scales more than they did members of the intermediate subgroups. The results of the peer ratings for Group I subjects were the following:

1. The high GCT subgroup and the RPT subgroup were rated significantly superior to the intermediate subgroups on personal cleanliness, military bearing, bunk and living quarters, general Navy performance, and leadership.

2. The RPT subgroup was rated as significantly superior to the other three subgroups on marching and maneuvers and physical drill with rifle.

3. The high GCT subgroup was rated as significantly superior to the other three subgroups on quickness to learn and response to orders.

4. The RPT subgroup was superior to the intermediate subgroups on care of clothing.

The results of Company Commanders' ratings of the Group I subjects were the following:

1. The high GCT subgroup was significantly superior to the other three subgroups on quickness to learn, response to orders, willingness to work, and general Navy performance.

2. The high GCT subgroup and the RPT subgroup showed superior ratings on care of clothing and military bearing.

The pattern of ratings for Group II was essentially the same as that obtained for Group I.

In general, it seems reasonable to conclude with Cofer that RPT aided recruits in adjusting to regular recruit training, but it is still difficult to determine what contributions the cognitive aspects of literacy training made to this adjustment.

Standlee (1954) studied three groups of recruits in his investigation:

1. RPT graduates: N = 611.

2. Marginal recruits with a GCT of 35 or lower, an NVCT of 38 or higher, and a LIT of 38 or higher: N = 2414.

3. Typical recruits with a GCT of 36 or higher: N = 1998.

The 5312 total subjects were originally selected by identifying all RPT graduates and marginal recruits processed during the period 10 May 1952 to 1 September 1952 and a random 10 percent of the typical recruits processed during the same period. Data on the fleet performance of these recruits were obtained by mailing out questionnaires to their Commanding Officers during July-August 1953. Most of the questionnaires were returned, 5023 of which were suitable for analysis--an effective response rate of almost 95 percent.

The RPT and marginal groups in comparison with the typical group were much more often assigned to general duty (unskilled work) and to duties that did not require as much reading skill. The two groups were less favored with promotions, had lower average performance evaluations by supervisors and Commanding Officers, and had a higher proportion of disciplinary actions and days lost due to misconduct or sickness. In the opinion of Division Leading Petty Officers, the two lower groups generally showed less promise for future advancement and profitable service to the Navy. The two lower groups, however, indicated more intention to reenlist than the "typical" group.

Overall, the two lower groups tended to resemble each other, and to be somewhat less effective in performance than the "typicals." Where differences between the two lower groups were found, they tended to show the marginal group to be superior to the RPT group. However, most of the differences between the three groups were small and none of the groups had characteristically unacceptable performance; the majority of members in all the groups were judged to be serving the Navy adequately.

Two other relevant investigations came out of the general contractual effort that sponsored the Cofer and Standlee studies. Standlee, Fattu, and Auble (1954b) investigated the quality of Navy technical writing by tabulating the frequency of words appearing in the Bluejackets' Manual (14th edition), This is Your Navy, All Hands, and the Naval Training Bulletin; and Fattu and Standlee (1954) applied two Flesch Readability formulas to the following publications: Bluejackets' Manual, Stewardsman (Manual), Fireman (Manual), Steward and Cook (Manual), Commissaryman (Manual), This is Your Navy, All Hands, and Naval Training Bulletin. It seems unlikely that these investigations represent the first attempts to systematically judge the quality of technical writing in the Navy, but these are the first that appear to be available in the official literature. The results of these investigations are predictable: Standlee et al. showed that even experienced writers and teachers use too many rare words and exclude too many common words when relying on their experience and common sense; Fattu and Standlee found that standard, essential Navy publications were, most probably, too difficult for Navy enlisted men--the average Flesch readability score of their sample was 61.7, which converts roughly to a ninth grade reading level.

With the end of the Korean War, sufficient manpower again became available to the Navy, and in 1957 the RPT program was discontinued in favor of a shift in research and administration emphasis toward problems connected with higher-level personnel. Subjectively, many persons in the Navy (including most of the RPT graduates) felt that RPT was well worth the time and effort; objectively, the evaluation findings were inconclusive.

Academic Remedial Training in the Present

The Navy's current excursion into literacy training appears to have been brought about by the Army's manpower supply problems during the Vietnam conflict. During the Vietnam buildup, the Army was forced to accept personnel who were classified as marginal by their scores on the Armed Forces Qualification Test (AFQT). Generally, these volunteers and draftees ranked in the 10th-30th percentile range on the AFQT (Project One Hundred Thousand, 1969). As a consequence of its manpower needs, the Army indicated to the Department of Defense that a coherent program for processing marginal personnel was needed if manpower requirements were to continue being met. Defense responded in October 1966, by establishing "Project 100,000," which was to help meet manpower supply problems by spreading marginal personnel throughout all three services. The project was also intended to provide training for these men so that they would be better prepared to return to civilian life if they chose to do so. Under Project 100,000, the Navy agreed to accept up to 15 percent of its enlistees from the marginal category and allow them to volunteer for the normal draft tour of 2 years. Accordingly, the Navy established RPT units at Great Lakes, Illinois, and San Diego, California, in February 1967 (Weeden, 1975). Almost immediately, the name of the program was changed to Academic Remedial Training (ART), and this is the name currently in use.

Despite the expiration of Project 100,000 in 1972, the ART program has continued in operation up to the present. Although projections have indicated that induction of marginal personnel will be unnecessary even under the current no-draft policy (e.g., Battelle, Brown, Kruzie, Marshall, Moll, Paskert, & Radovich, 1973), ART may be continued because of the rising national consciousness concerning equal employment opportunity and/or because of (anticipated) improvements in the national economy.

A student may currently be selected for ART if he averages between 3.0 and 5.5 grade levels on the Vocabulary and Comprehension Sections of the Gates-MacGinitie Reading Test. Originally, the Navy Life series was once more dusted off and reedited for the program; however, use of these materials was largely discontinued in 1969. As described by Weeden (1975), the course of instruction may last a maximum of 8 weeks and is broken into two phases: a diagnostic phase and a therapeutic phase. The diagnostic phase provides for initial assignment of the student to one of three courses:

1. Course Mike--a phonics course emphasizing word attack skills.

2. Course Oscar--a reading course beginning at the 3.0 grade level.

3. Course Victor--a reading course beginning at the 4.0 grade level.

The therapeutic phase consists of two mutually reinforcing areas: word attack and reading abilities. The word attack area concerns phonics, vocabulary development, and word knowledge (root words, inflections, prefixes, synonyms, etc.). The reading abilities area concerns reading practice, comprehension skills, and study skills.

Notably, ART is integrated only into recruit training as were RPT and the Special Training Program before it. The proportion of illiterates in the United States in 1960 was estimated to be only 2.4 percent (Harman, 1970), and, with current screening practices, it seems unlikely that there is an appreciable number of illiterates assigned to Navy billets. However, a new concern that keys on the concept of functional literacy is beginning to appear. It may well be that, despite the universal attainment of literacy by Navy personnel some men (and women) may fail to perform because they do not read well enough to meet the requirements of their jobs; there may, in effect, exist a literacy gap. This possibility was first raised by the Fattu and Standlee report of 1954; however, two reports by Carver (1973a, 1973b) and current work by Duffy at the Navy Personnel Research and Development Center seem to support the current concern over a literacy gap. If such a gap exists it may become necessary to extend literacy training from the recruit commands to the fleet. The nature and location of programs devised to meet the literacy gap would represent very new directions in the history of literacy training in the Navy.

Final Comment

From the preceding survey of literacy training, it is apparent that problems of literacy are not a recent discovery; they have been a concern of the Navy throughout its history. This concern has been motivated both by a need for proficient job performance and by an interest in the general welfare of Navy personnel. In the development of literacy programs, the Navy has, at one time or another, been sensitive to at least the following six issues:

1. Level of Literacy. Efforts have been made to identify levels of literacy and to tailor instruction to the level required of a trainee. The idea of functional literacy as a consideration in determining what level of literacy a trainee ought to achieve in order to perform a specific job has also appeared.

2. Cause of Literacy Problems. Adjustments in Navy programs have been made to identify speakers from nonstandard English linguistic communities and to adjust instruction based on the language habits likely to be fostered by those communities.

3. Evaluation. Both narrow (Are program objectives being met?) and broad (Is Fleet performance being enhanced?) issues have been considered in evaluating Navy literacy training.

4. Decoding. Issues of relating orthography to acoustic representation so that learners may identify information within their linguistic experience have been addressed.

5. Affect. Learner's attitudes toward themselves and toward literacy have been taken into account.

6. Active Participation. Most military training attempts to involve as much "hands-on" experience as possible and this approach appears in Navy literacy training, particularly with the emphasis on the Navy Life workbooks.

On the other hand, the Navy's interest in literacy has been largely intermittant. From the evangelical concern of the chaplains to the current interest in ART, lack of literacy skills on the part of Navy personnel is always given full opportunity to handicap the Navy's operational effectiveness before anything is done about it. Programs that are devised in response to literacy problems usually prove to be relatively sophisticated as the Special Training Program and RPT curriculums demonstrate, but these programs seem to appear only in response to wars or other national emergencies. The outlook for the current ART program cannot be very bright. ART is remarkable for continuing as long as it has after the demise of Project 100,000, but it seems reasonable to anticipate waning interest in ART on the part of the Navy unless (1) problems arising from the all-volunteer military suddenly become more severe, (2) manpower supply, for whatever reason, becomes more constrained, or (3) the national priorities for equal employment opportunities are sustained.

The primary hope for continued interest in literacy training seems to rest on the concept of functional literacy in which the concern is not so much with illiteracy as with what literacy is necessary for specific Navy jobs (e.g., Duffy, Carter, Fletcher, & Aiken, 1975). In this respect, it seems likely that concerns of literacy will be modified in two ways. First, they will become more narrow in that finding and applying technical information will be emphasized rather than more general skills of reading and writing. Second, these concerns will become more broad in that considerations of media, such as the organization and quality of technical manuals, the availability and capacity of systems for computer-based information retrieval, and the comprehensibility and usability of graphic and pictorial information, will be integrated with more conventional literacy considerations. These are fairly novel directives for literacy training in the Navy, but they are also timely. It is difficult to come away from an historical survey of literacy training without an appreciation for the continuing increases in the amount of information that must be available for successful performance of Navy jobs. Classically exponential, these increases were at first gradual but are now accelerating at an impressive rate. The importance, and even urgency, for systematic concern with literacy is hard to escape, and it is to be hoped that literacy training will receive continuing support from all the military services.

References

Battelle, R. B., Brown, H. D., Kruzie, P. G., Marshall, T. H., Moll, K. D., Paskert, P. F., & Radovic, M. Analysis of some potential manpower policies for the all-volunteer Navy. Menlo Park, CA: Stanford Research Institute, June 1973.

Bloomfield, L., & Barnhart, C. L. Let's read: A linguistic approach. Detroit, MI: Wayne State University Press, 1961.

Burr, H. L. Education in the early Navy. Doctoral dissertation, Temple University. Philadelphia, PA: Published by the author, 1939.

Carver, R. P. Measuring the reading ability levels of Navy personnel. Washington, D. C.: American Institutes for Research, October 1973. (a)

Carver, R. P. Measuring the reading difficulty levels of Navy training manuals. Washington, D. C.: American Institutes for Research, October 1973. (b)

Cofer, C. N. Adjustment to recruit training: A study of the effects of recruit preparatory training (Technical Bulletin 54-22). Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, December 1954.

Couch, P. E. Educational emphasis in civilian conservation camps of the Seventh Corps area. Unpublished doctoral dissertation, Indiana University, 1944.

Cummings, D. E. Enlisted training in the Navy. United States Naval Institute Proceedings, 1929, 55, 878-886.

Curriculum for Recruit Preparatory Training. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, April 1953.

Drury, C. M. The history of the Chaplain Corps United States Navy. Washington, D. C.: Government Printing Office, 1949.

Duffy, T. M., Carter, J. C., Fletcher, J. D., & Aiken, E. G. *Language skills: A research and development plan for the Naval service* (Special Report 76-3). San Diego, CA: Navy Personnel Research and Development Center, October 1975.

Fattu, N. A., Fay, L. C., D'Amico, L. A., & Standlee, L. S. *Observations on Navy literacy training* (Memorandum Report). Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, March 1954.

Fattu, N. A., Mech, E., & Standlee, L. S. *A review of literacy training programs in the armed services during World War II* (Technical Bulletin 53-4). Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, December 1953.

Fattu, N. A., & Standlee, L. S. *Analysis of reading difficulty of selected Navy materials* (Technical Bulletin 54-3). Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, March 1954.

Fullam, W. F. *The Petty Officer's Drill Book, United States Navy*. Annapolis, MD: United States Naval Institute, 1902.

Fullam, W. F. *The Recruit's Handy Book*. Annapolis, MD: United States Naval Institute, 1902.

Ginzberg, E., & Bray, D. W. *The uneducated*. New York, NY: Columbia University Press, 1953.

Hagen, E. P., & Thorndike, R. L. *A study of the World War II Navy careers of illiterates sent through literacy training* (Res. Rep., Contract NONR-644(00)). Washington, D. C.: Bureau of Naval Personnel, Classification and Survey Research Branch, April 1953.

Harman, D. Illiteracy: An overview. *Harvard Educational Review*, 1970, *40*, 226-243.

Harrod, F. S. *Enlisted men in the United States Navy 1899-1939*. Unpublished doctoral dissertation, Northwestern University, June 1973.

Heath, S. R., Jr. The teaching of reading in a military setting. *The Training Bulletin*, 1946, *43*, 150-156.

Hodges, C. I. *Special training programs* (Unpublished Report). Washington, D. C.: Bureau of Naval Personnel, November 1964.

Langley, H. D. *Social reform in the United States Navy, 1798-1862*. Urbana, ILL: University of Illinois Press, 1967.

Luce, S. B. The manning of a Navy and mercantile marine. *United States Naval Institute Proceedings*, 1874, *1*, 17-37.

Luce, S. B. Naval training. *United States Naval Institute Proceedings*, 1890, *16*, 367-430.

McLean, R. Bluejacket's Manual. Annapolis, MD: United States Naval Institute, 1902.

Niblack, A. P. The enlistment, training, and organization of crews for our new ships. United States Naval Institute Proceedings, 1891, 17, 3-49.

Potter, F. H. A school for bluejackets. Outlook, 1913, 114, 694-705.

Potter, F. H. A repair shop for men. Outlook, 1918, 120, 539-540.

Project one hundred thousand: Characteristics and performance of "new standards" men. Washington, D. C.: Department of Defense, Manpower and Reserve Affairs, March 1969.

Ross, C. S. Literacy training in the Navy. School and Society, 1946, 63, 203-204.

Special training program (Unpublished Report). Washington, D. C.: Bureau of Naval Personnel, 1951.

Standlee, L. S. A follow-up comparison of three groups of Navy enlisted men: Marginal-and-illiterate, marginal-but-literate, and typical recruits (Technical Bulletin 54-20). Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, July 1954.

Standlee, L. S., Fattu, N. A., & Auble, J. D. Data on daily performance of marginal-and-illiterate, marginal-but-literate, and typical recruits from three naval training centers (Memorandum Report). Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, July 1954. (a)

Standlee, L. S., Fattu, N. A., & Auble, J. D. Frequency index of words appearing in four Navy publications (Technical Bulletin 54-2). Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, January 1954. (b)

Weeden, E. T. Academic remedial training (Unpublished Report). San Diego, CA: Naval Training Center, January 1975.

LITERACY TRAINING IN THE NAVY

T. M. Duffy

Literacy research in the Navy today has had two major objectives. First, there have been efforts to characterize the literacy "problem" in the Navy and thereby provide a data base for evaluating alternative approaches to insuring that a literacy deficit at any level does not hamper the effectiveness of the service. The full range of reading levels has been examined and basic demographic data on reading and its relationship to other abilities and to background characteristics have been provided. This research has also involved an examination of the relationship of reading skill to job performance. The second major line of research has focused on the Navy's current approach to deficient literacy skills--reading training programs for personnel with marginal reading ability. This research has involved evaluations of the effectiveness of the training programs, tests of new instructional methods, and procedures for predicting the successful reading training student. Before examining this literacy research in detail, I would like to briefly characterize the Navy's objectives vis-a-vis literacy and the options available in meeting these objectives. This characterization will provide both an understanding of the Navy's current re-emphasis on literacy and a framework for many of the questions addressed in the research efforts.

The Marine Corps currently provides basic skill training in reading at both of their recruit training centers for all recruits reading below the 4.5 reading grade level (RGL). For the Navy, reading training is provided at the three Recruit Training Centers for all recruits having a tested RGL between 3.0 and 5.5. In Fiscal Year (FY) 1975, an estimated 5.1 percent of the Navy recruit population, or 4,500 men, had reading skills in this range. Recruits below a 4.0 RGL are considered poor candidates for the short-term reading training and are typically discharged from the service, while recruits above a 5.5 RGL are considered to have adequate reading skills. Although the reading programs focus on personnel with marginal capabilities, the objectives of the programs and the Service's objectives in literacy are more generally stated. These objectives are: (1) to provide a level of literacy skill to all personnel so as to insure Fleet effectiveness and Fleet safety, and (2) to provide the literacy skills necessary for equal opportunity in attaining upward mobility and a successful career (Stewart, 1974; Academic Remedial Training, 1975). Recent studies by Carver (1973b, 1973c) have raised the question as to whether either of these objectives is attainable within the structure of the current reading training programs. Carver compared the reading requirements of a sample of 20 Navy training manuals to the reading ability of a sample of 271 recruits. The manuals required, on the average, a 14th grade reading ability and ranged from 9th to 20th grade as measured by several readability formulas. In contrast, the average reading ability of the recruits was only 9th grade.

Carver's results indicated a major gap between the reading requirements and the reading abilities in the Navy. Additionally, they suggest that literacy training to the 5.5 RGL is less than adequate for meeting the literacy demands in the service. Regarding the objectives of literacy training, a 5.5 RGL would seem to neither insure Fleet safety and effectiveness nor to provide the opportunity for upward mobility. To meet these objectives will clearly require a reduction in the difficulty of manuals. In addition, personnel decisions with regard to literacy will be required. The options here include: (1) an expanded reading training program, (2) assignment to job areas with reading requirements matching reading ability, (3) selection into the service of only those men having adequate reading skills, or (4) some combination of the above. (See Duffy, Fletcher, Carter & Aiken, 1975; and Sticht, 1975b for further discussion of these options.)

While discussion of these options was underway,[1] few basic data were available upon which to base decisions. The distribution of reading levels in the Navy was unknown,[2] and thus the number of men affected and the cost involved in any of the options could not be determined. The literacy requirements in the various occupational areas were similarly unknown and thus a determination of which men an expanded literacy program should address and where in the career such a program should be located could not be made. Indeed, the very need for literacy training could not be assessed. Two arguments against the need for literacy training of any sort have been advanced. First, the move of the Service toward accepting only high school graduates has been proposed as an action which will also eliminate low-literate personnel. Second, it has been suggested that most equipment and vocabulary knowledge is acquired on the job and, therefore, with job experience a man will be able to read the necessary manuals. The argument continues that the inexperienced man, regardless of his reading ability, receives his instruction orally from his supervisor and, thus, his job does not require reading. An examination of the relationship of reading ability to education and to job performance would provide the data necessary to assess these arguments.

Reading Skill and Its Relationship to Other Personnel Characteristics

In 1974 a group of us at the Center began a research effort to characterize reading skills in the Navy and thus provide a data base to answer some of the questions described above. The program was a joint effort with the Naval Training Center and Recruit Training Command in San Diego and had as an additional goal the early identification of all recruits eligible for the San Diego reading program. We have now collected reading and other ability data as well as performance data on approximately 30,000 recruits and on 1500 men receiving occupational ("A" School) training.

---

[1] Chief of Naval Operations (OP-099) speedletter 991B/550 of 3/13/74, Readability Level of Publications and Adult Basic Skill Training.

[2] The distribution and variability of reading scores was not available from the Carver (1973b) data.

Method

Recruit Sample and Reading Test. All recruits entering recruit training at San Diego from May 1974 to May 1975 were administered the Gates-MacGinitie Reading Test, Survey D (Gates & MacGinitie, 1965). Only the vocabulary and comprehension subtests were administered. The vocabulary subtest consists of 50 multiple choice items and requires synonym matching. The comprehension subtest involves a modified cloze procedure (Taylor, 1957) in which words are deleted from paragraphs and the subject selects the proper word from several alternatives. The test is directed at 4th and 6th grade readers, but is empirically normed on students in the 3rd through 9th grades. The test was administerd during the first week of training, a processing week during which clothes, etc., are issued and standard Navy tests are administered. Testing was done in groups of 50 to 70 men (by recruit companies).

During our testing period, 32,890 men were processed at San Diego. Of this total, 96 percent, or 31,540 men, were administered reading tests. In examining the relationship of reading scores to standard Navy test scores, the sample was reduced to 28,542 or 86.7 percent of the total input. The reduction in sample size is due to a variety of factors, but of primary importance were the discharge or hospitalization of men at the time of testing and the incorrect recording of identifiers at the time of testing, which resulted in an inability to match a man's records at the time of data analysis. The data analyses did not always involve the entire sample, but in all cases the sample used involved all of the men processed in a specified time frame.

"A" School Sample and Reading Test. Men from the "A" Schools at the San Diego Training Center were represented in this sample. The "A" School is the first specialized training a man receives after recruit training. Approximately 70 percent of Navy recruits attend one of the 50 or more "A" Schools. The San Diego schools in our sample ranged from clerical to electronics training.[3] Testing occurred in the classroom at each school during class. Men were tested from April to September 1975. The sample was doubled in four of the short-duration schools so as to have equivalent time frames for the samples. At present we only have data available on six schools, involving a total of 564 men.

The Nelson-Denny Reading Test (Nelson & Denny, 1960) was administered to the "A" School sample. Again, the vocabulary and comprehension subsections were used. The vocabulary test is similar to the Gates-MacGinitie. The comprehension test involves reading paragraphs and answering fact and inferential questions about the content. This test is directed at the 9th to 16th grade reader and is empirically normed on students in grades 4 through 16.

---

[3] The schools were: Ship's Serviceman, Hull Technician, Mess Specialist, Interior Communications, Electricians Mate, Basic Electronics & Electricity, Quartermaster, Machinery Repair, Data Processing, Signalman, and Radioman.

Other Test and Personnel Data. All of the "A" School sample and 1294 recruits were also tested on the Navy Pattern Matching (PM) test of nonverbal ability. The PM was derived from the Raven's Standard Progressive Matrices Test (Raven, 1958), which is generally considered to be a culture-free test of nonverbal intellectual ability (Jensen, 1972; Carver, 1973a; Spearman, 1946). The derivation of the PM involved an item analyses of the Progressive Matrices in which low frequency item alternatives and non-discriminating items were eliminated. The test involves 38 multiple-choice items with five alternatives per item. Norming on a sample of 500 Navy recruits yielded a mean of 28.8 and standard deviation of 5.3.

For both the recruit and "A" School samples, we obtained all of the test and background data collected as a normal part of processing into the service. These data include:

1. General Classification Test (GCT)--a test of general ability involving verbal analogy and sentence completion items.

2. Arithmetic Reasoning Test (ARI)--a test of mathematical ability involving word problems.

3. Mechanical Ability (MECH)--a low-verbal test of knowledge of mechanical principles.

4. Clerical Test (CLER)--a speeded digit search test requiring no verbal skill.

5. Electronics Test (ETST)--a verbal test of electronics aptitude.

6. Shop Practices Test (SHOP)--a test of tool knowledge requiring the matching of a picture of a tool to verbal descriptions of uses.

7. Armed Forces Qualification Test (AFQT)--a score derived from the GCT, ARI, MECH, and CLER by their regression on the former AFQT test of general ability, which was administered throughout the Armed Forces.

8. Years of Education--self reported.

9. Race--obtained only for the recruit sample.

Results

Reading Levels. Figure 1 presents the distribution of reading levels in the recruit sample. The mean score of 9.8 compares favorably with the mean of 9.3 obtained by Carver (1973c). However, because of the skewness, the median score of 10.7 is more reflective of the general reading ability. The skewness of the distribution reflects the limitation of the Gates-MacGinitie test (which has a maximum score of 12.0 RGL) for this sample.

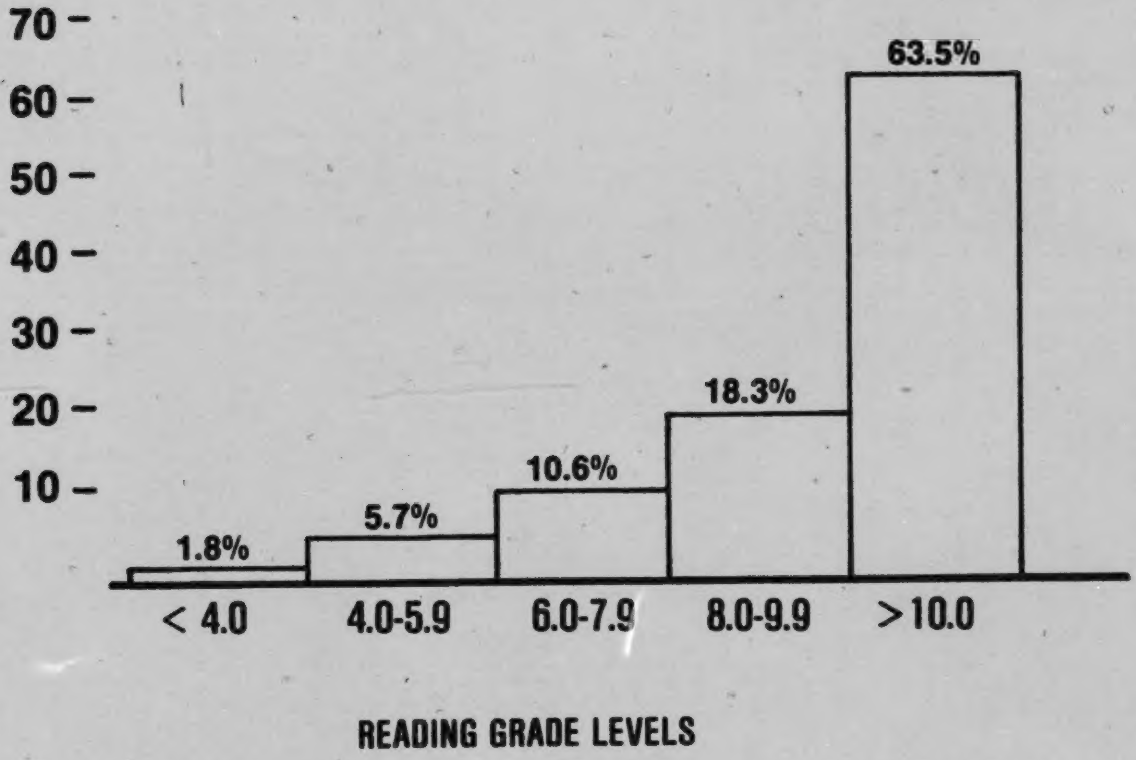


Figure 1. Distribution of reading levels of recruits at San Diego Recruit Training Command (N = 24,729).

The median RGL of 10.7 is generally reflective of the education level in our sample, where 70.9 percent left school after either the 11th or 12th grade. However, as indicated in Figure 1, a significant number of recruits have reading skills below the high school level. One method of determining the significance of these reading levels is to compare the reading ability of the men to the reading demands they face in the Navy. In this regard we analyzed the readability of those manuals and tests which most recruits are expected to read and comprehend. The materials were analyzed using the FORCAST formula (Caylor, Sticht, Fox, & Ford, 1973), which was developed on Army men and materials and relates the proportions of one-syllable words to the reading ability required for comprehension. The materials assessed and their reading difficulty were: The Bluejackets' Manual (11.5 RGL), which is the basic manual in boot camp; the Airman (10.5), Fireman (10.2) and Seaman (10.2) rate training manuals used during and right after completing boot camp; the General Classification Test (10.8) and Arithmetic Reasoning Test (9.0), which are administered in the process of classifying recruits. A comparison of these readability levels with the reading skills of recruits indicates a reasonable match between the average difficulty level of materials and the average skill level. However, the materials we assessed are materials all Navy recruits are expected to read and comprehend. Viewed in this way, the data indicate that approximately half of the recruits have a reading deficiency relative to their job-reading demands. Further, the deficiency amounts to a grade level deficit of 2.5 or more years for 18 percent of the recruits.

In the development of readability formulas, the baseline or norming material is read only once and comprehension (verbatim or inferential) of all of the material is assessed. Thus, in interpreting the difficulty level of material as assessed by a readability formula there is an implicit assumption that the manual is read only once and that all of the material is equally relevant. This is clearly not the case. The material is frequently read several times and is supplemented by oral instruction. Thus the apparent "literacy gap" as indicated by a comparison of reading levels and readability is most likely larger than the actual "literacy gap."

Regardless of the exactness of our determination of reading requirements, our ability data indicate that significant portions of Navy personnel will have difficulty comprehending typical Navy written materials. If it is necessary for these low ability men to comprehend general Navy printed materials, instructions, safety information, and perhaps technical information, then their effectiveness in the service should be reduced. (Evidence to be presented later indicates that reading skill is indeed related to performance.) For example, 1.8 percent of the recruits, or 1600 men (based on FY 1974 recruit figures), read below the 4th grade level. These men cannot decode the printed word and by national standards are considered functionally illiterate. Clearly, these men will not be able to read any job or safety material. The 4th to 8th grade reader is able to decode print but is deficient in vocabulary skill, in the ability to infer from and interrelate printed information, and in reading speed. In our recruit sample 18 percent, or a projected 15,800 men, entered the Navy last year with reading skills below the 8th grade level. These men may be expected to comprehend most Navy material if they can reread the material and are assisted by other personnel. However, the efficiency of their job performance will be impaired and at a marginally acceptable level. If these men are required to act in an emergency in any way that requires the use of printed material, they may well prove a hazard to effective Navy operations.

A minimum of 70 percent of recruits go on to "A" School training, so there is considerable overlap in the recruit and "A" School populations. While the primary purpose of the "A" School assessment was to examine the role of reading in school performance, the reading levels are of interest themselves since they provide a more complete picture of reading skills in the Navy and the options for assuring adequate skill levels. For example, the presence or absence of substantial numbers of low skill readers in the "A" Schools has implications for classification policies as well as the location and nature of reading training programs.

The Nelson-Denny reading test was used in assessing the "A" School sample, since it provides norms for higher levels of reading skill.[4] The mean and median RGLs in this sample were 11.1 and 11.2, respectively. Even though the "A" School personnel are selected on the basis of ability, we find that 9.6 percent of the sample had reading skills below the 8th grade level. Table 1 presents a more detailed examination of the reading levels in each of six "A" Schools as well as an indication of the reading requirements (readability) in those schools.

The readability score is for the rate training manual, which is studied for advancement to the 3rd and 2nd class rates. The manual is used in the "A" School classroom and the rate exam is typically taken within four months of completing the school. The data in Table 1 indicate that for all but the Ship's Serviceman, the mean reading ability of the men approximates and in some cases exceeds the mean reading difficulty of the manual. However, in three of the six schools a significant proportion of the men are low-ability readers and may be expected to face reading difficulties. Even in those ratings where the proportion of low ability readers is small, the actual number of men, several hundred per rating based on FY 1974 manning figures, is considerable.

---

[4] Data on 188 men who appeared in both samples indicate that the Nelson-Denny and Gates-MacGinitie tests yield similar data. The median reading ability for these men was 11.5 on the Gates as compared to 12.5 on the Nelson-Denny. The correlation of the two test scores was .65, which is a strong correlation when one considers that half of the men in the sample fell between the 11.5 and 12.0 RGL on the Gates-MacGinities test.

Table 1

Reading Grade Levels (RGL) of Men and Reading Grade Level Requirements For Manuals in Each of Six Navy Occupational Training "A" Schools

	Men		Manuals
School	X RGL[a]	% Below 8.0 RGL	X RGL[b]
Hull Maintenance Technician	10.2	19.1%	10.7
Ship's Serviceman	10.4	18.1%	12.9
Mess Specialist	10.6	12.0%	10.2
Quartermaster	11.7	2.9%	10.9
Interior Communications	11.9	5.1%	12.6
Electricians Mate	12.0	2.3%	12.7

[a] RGL for men is based on Nelson-Denny reading test performance. Means and medians were comparable in each school.

[b] Readability scores are taken from Biersner (1975) and are based on the application of the Flesch Reading Ease formula normed on Navy men and manuals (Kincaid et al., 1975).

All of our assessment data indicate that the options available in approaching literacy deficits in the Navy will affect significant numbers of personnel. For example, if a 5.5 RGL, the current reading level deemed necessary in the Navy, were required for admission into the service, 9 percent, or 7800, of the FY 1974 recruit population would have been rejected. Eighteen percent of the recruits, or 15,800 annually, would be eligible for a literacy program expanded to an 8.0 RGL terminal criterion, a criterion which only approaches the initial reading demands in the service. The same number of men would be affected by the implementation of limited duty assignments for below 8th grade readers. The costs involved and the number of men affected by implementation of any of these options demand that the options be carefully considered.

Our findings with regard to literacy skill levels in the Navy indicate that one or more of the above options must be implemented to assure effective performance levels in the service. These options, however, must necessarily focus on the very low literate man in order to be manageable and within reasonable cost figures. The distribution of reading skills in the "A" Schools suggests an additional option which would address personnel with less severe literacy deficits. This option involves a literacy training program for personnel deficient in reading skills but qualified for "A" School. The program would have as a criterion the ability demands of each school and would provide functional reading training preparatory to, or in conjunction with, school training. The major benefit of this program is that it would provide the opportunities of upward mobility and career flexibility for men who, while eligible for "A" School training, are likely to perform at a marginal level. Additionally, the program would involve fewer personnel and hence would be more manageable than a broad-based literacy program. Finally, the program would address specifiable Navy reading requirements. HumRRO, under contract to the Center, is currently evaluating the cost effectiveness and management considerations involved in this and other literacy training options.

Education. One of the suggestions discussed earlier was that the move toward accepting only high school graduates would reduce or eliminate the marginally literate personnel. Our data indicate, to the contrary, that a high school diploma bears little relationship to reading skills of personnel currently in the service. As can be seen in Figure 2, the median reading level for men with at least a high school diploma is less than 1 RGL higher than that of the non-high school graduates (the medians are 10.9 and 10.2 respectively). While proportionately fewer high school graduates fall below an 8th grade RGL, this proportion (14.9 percent), nonetheless, indicates that the selection of only high school graduates would still produce a significant number of marginally skilled individuals.

Further analysis on a sub-sample of 19,000 recruits indicates that years of education, ranging from 8th to 16th grade, only correlates .13 with reading ability. Of all of the test and background information for which we have data, years of education shows the least relationship to reading. This finding may be contrasted with previous research which has found years of education to be the most valid predictor of attrition from the service (Plag & Hardacore, 1964) and delinquency (Gunderson & Ballard, undated). One might conclude from these data that a high school diploma indicates a person's willingness to conform to the rules of society rather than his ability. However, a more likely explanation is that over recent years the predictive power of years of education has been reduced due to the increasing proportion of high school graduates entering the Navy. Thus, the small correlation we obtained is likely due to the reduced variance in years of education as well as the reduction in the number of personnel with very low education levels and reading skills rather than to any change in the basic relationship between reading and education. Indeed, recent evaluations of adult reading skills in the United States indicate that education, in this population, is predictive of reading ability. In studies by Northcutt, Selz, Shelton, Nyer, Hickok, and Humble (1975) and Young and Jamison (1975) a reading test was administered to a large, representative sample of U. S. adults. In both studies, education level was the demographic variable which most strongly predicted reading skill. Thus, the Navy's current policy in rejecting most applicants with little education also tends to eliminate the poor readers. However, the data in Figure 2 indicate that any further restrictions in educational requirements would have little effect on the reading skill levels in the Navy.

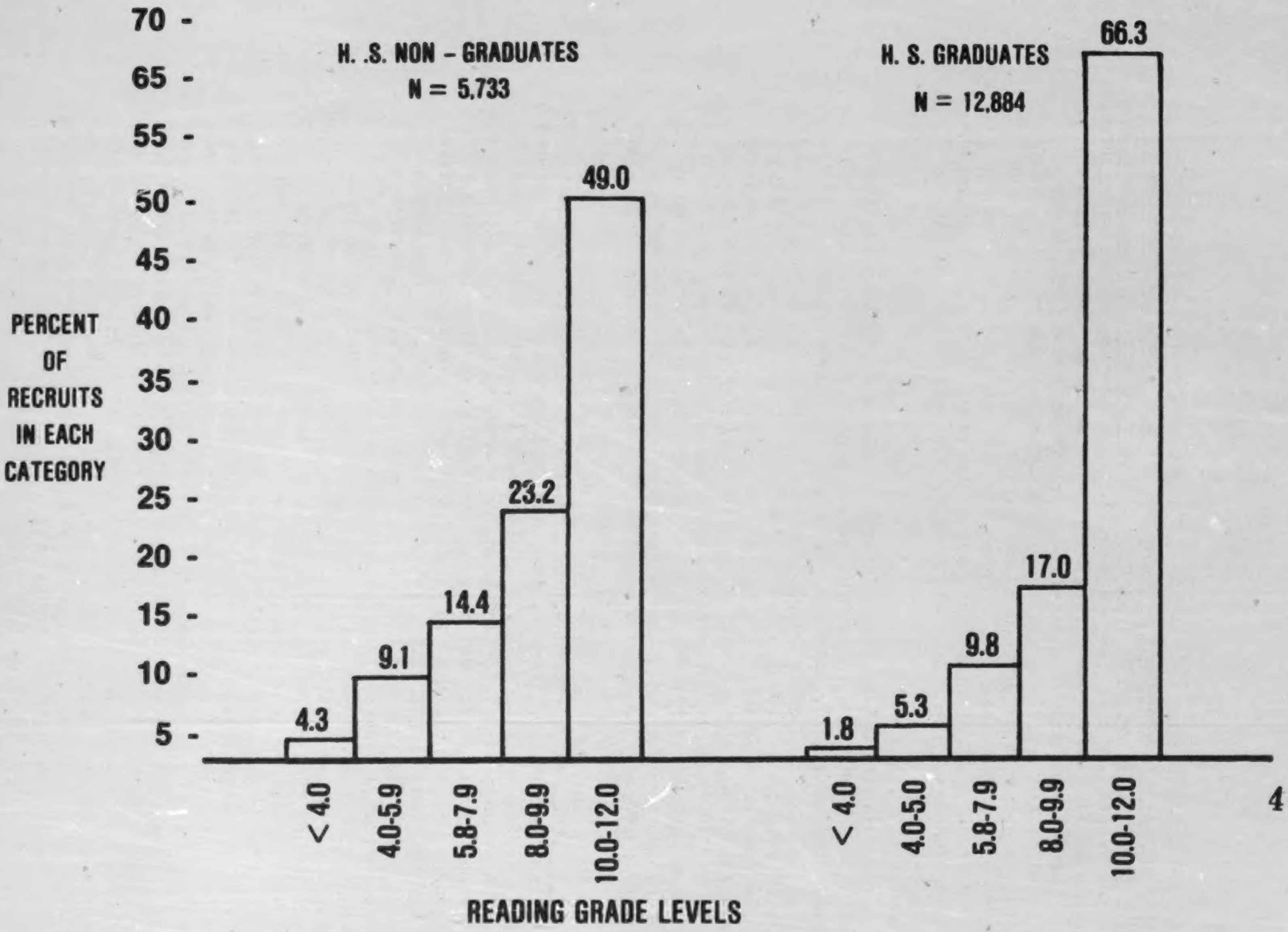


Figure 2. Reading levels for high school graduates and non-graduates at RTC San Diego,   May-Oct 1974 (N = 7138).

Race. Figure 3 presents the distribution of reading scores for each of the three major race categories at the San Diego Recruit Training Center (RTC). Aproximately 1.5 RGLs separate the median scores of the groups, with Caucasians having the highest median (11.0) and Malaysians the lowest (7.6). The same ranking of races is obtained when one compares the proportion of men reading below the 8th grade level. However, in terms of absolute numbers the men with lower reading skills at the San Diego RTC are primarily Caucasians and Malaysians.

The race data suggest a major cause of many of the reading problems for men at the San Diego RTC. The Malaysians, while representing only 8.6 percent of the recruit population, constitute 28.0 percent of the population of below 8th grade readers. The Malaysians also constitute a group in which English is a second language for most men. Although English is taught beginning in the 1st grade in the Malaysian countries, its use is restricted almost entirely to the school and is limited to "formal English." Thus, a deficiency in functional reading skills may be expected.[5]

The relationship of reading skill and racial category indicate that any policy of selection or classification on the basis of reading skills would have definite affects on the racial distribution in the Navy. The data also indicate that any reading training program will have to have the capability of dealing with skill deficiencies characterizing personnel with English as a second language as well as personnel with English as a native language but who have failed to develop adequate reading skills.

Ability Tests. One of our primary interests in the ability test data was to determine the degree to which reading ability could be predicted from performance on standard Navy tests. Our testing program was limited to San Diego and was due to terminate in March of 1974. A continued tracking of reading levels in the Navy, continued estimation of the number of low-ability readers, and estimation of reading levels at the other RTCs would be possible if a strong relationship were obtained between reading and the Basic Battery Tests.

---

[5] The race data are not necessarily representative of the Orlando and Great Lakes RTCs since the large proportion of Malaysians at San Diego is due to its geographical location. However, considerations of English as a second language may apply to the segment of Spanish speaking people found at those training commands.

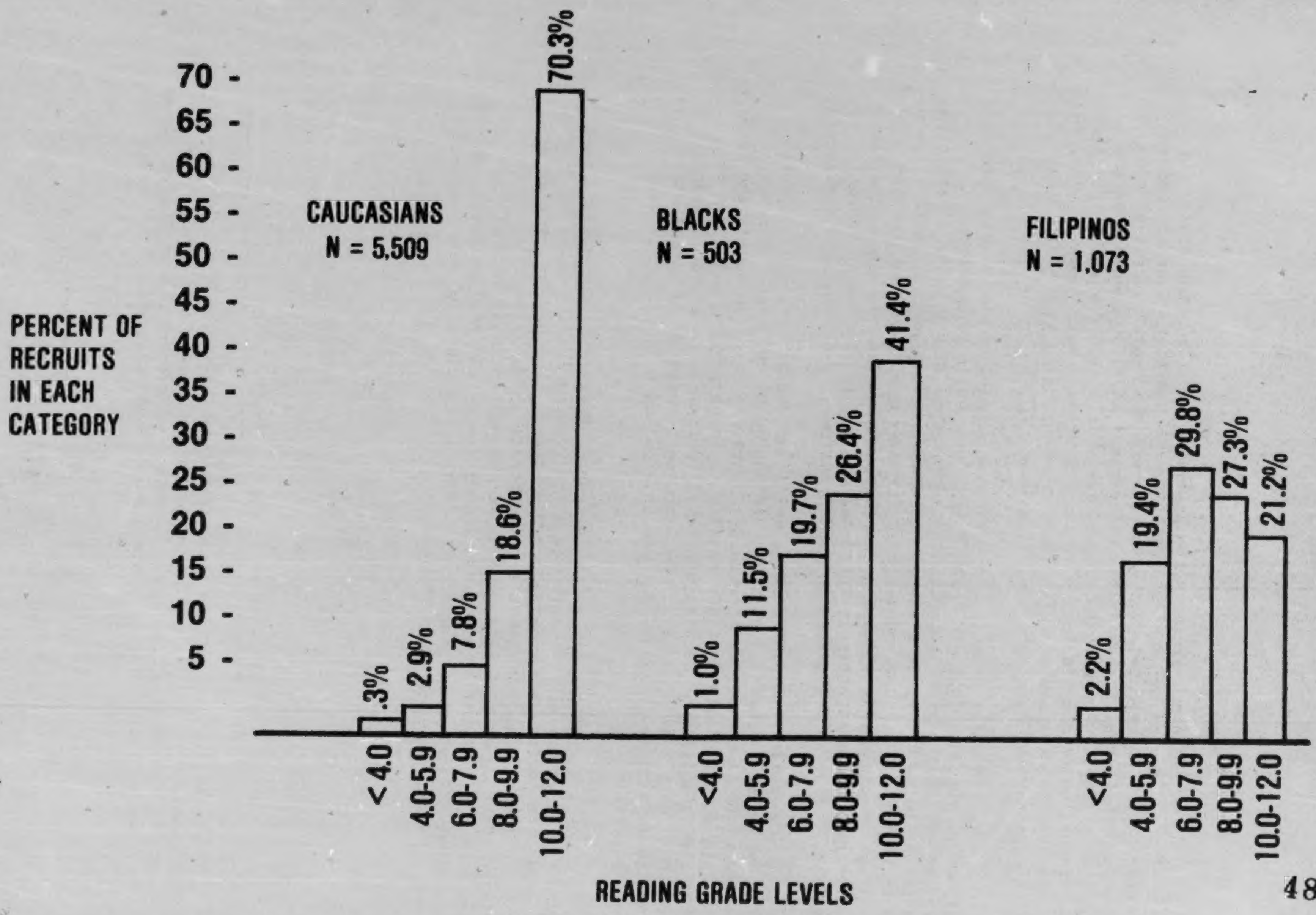


Figure 3. Reading levels in major race categories at RTC, San Diego, May-Aug 1974.

Data from the first 7135 recruits taking the reading test were used in the initial analysis. All Basic Test Battery scores were entered as prediction variables into a stepwise multiple regression analysis with the Gates-MacGinitie reading score as the dependent variable. The GCT, a highly verbal general ability test, was found to be the best predictor of reading, $r = .73$ and standard error = 1.36 RGL. The clerical test entered next into the regression equation but added only .006 points to the multiple correlation. Thus the prediction of reading is most efficiently based on GCT alone. These findings are similar to the reading--general aptitude correlations of .68 and .82 obtained in the Army and Air Force respectively (Caylor et al., 1973; Madden & Tupes, 1966). The degree of relationship is to be expected since their general aptitude tests, like the Navy's GCT, tend to be highly verbal tests.

The empirically determined RGL as well as the predicted RGL for each level of GCT is presented in Figure 4. The empirical relationship is linear in the midranges, but both a basement and a ceiling effect are evident. The ceiling effect is due to the limitation of the reading test, while the basement effect is likely due to a lack of sensitivity of the GCT at the low end since the GCT was designed to predict school success rather than to discriminate between low ability personnel. The validity of the regression of GCT onto reading was checked using a sample of 4517 recruits. This sample consisted of all recruits entering RTC San Diego during the two-month period following the initial sample. The cross-validity coefficient for reading and GCT dropped to .64 with a standard error of 1.6. The reason for the magnitude of this decrease in relationship is unclear, but it would seem to be partly a chance fluctuation since additional analyses involving GCT and reading in other recruit samples have yielded correlation ranging from .69 to .80. In any case, comparing the standard error of estimate obtained on the initial and cross-validation samples indicates that the decrease in correlation should not seriously affect the accuracy in predicting the general levels of reading ability in the service.

The results of these analyses indicate that GCT can be used to track changes in reading skills in the Navy's accessions. In addition, GCT can be used to obtain an approximation of proportion of recruits falling into gross categories of reading skill. However, GCT cannot be used to discriminate between low ability readers. At best, the results indicate that 95 percent of those recruits reading below the 6th grade level will have a GCT which is at least one half of a standard deviation below the mean (less than a 45 GCT). In the future we will be comparing this regression equation with one derived from the Nelson-Denny test in our "A" School sample.

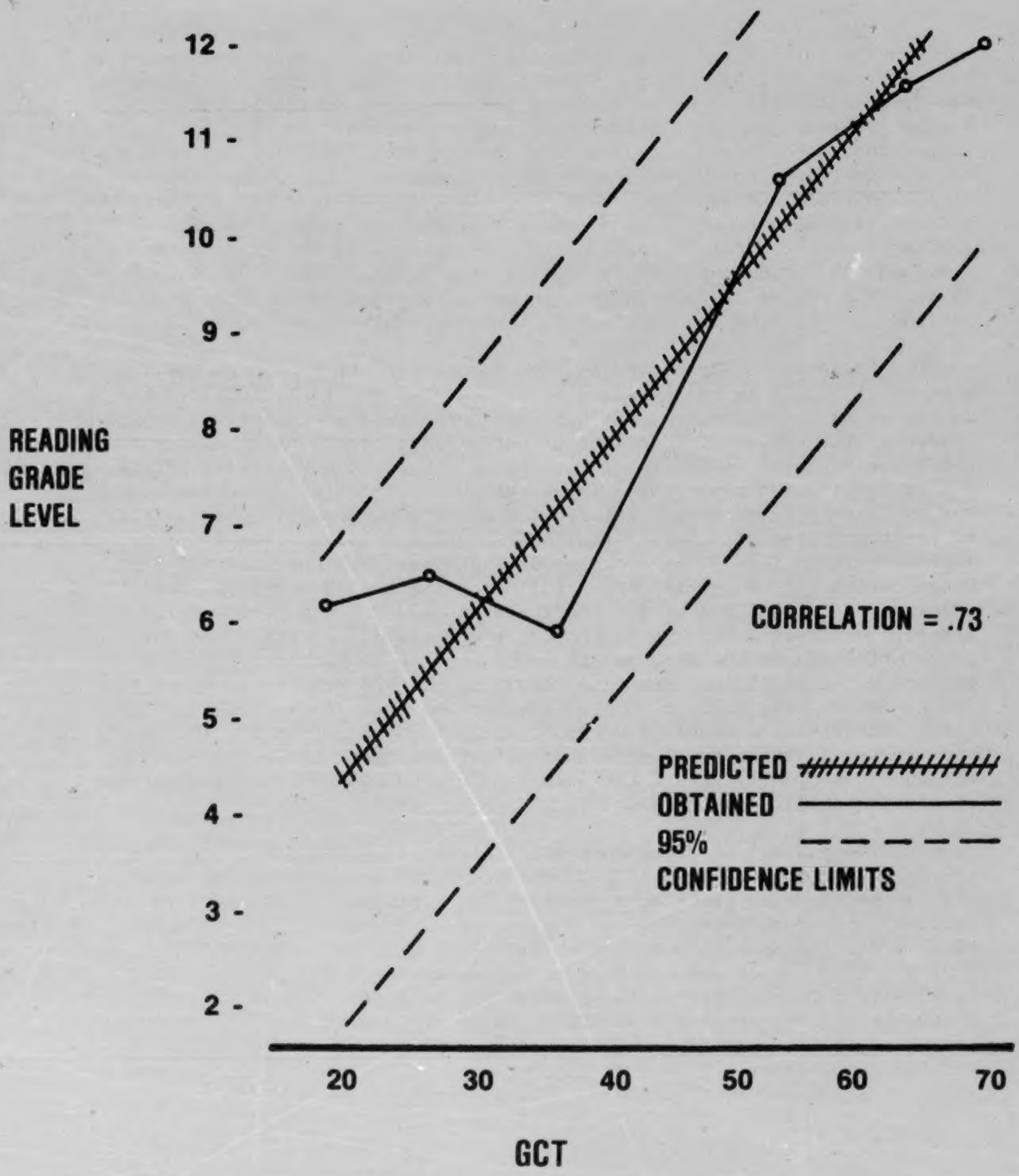


Figure 4. Relationship between GCT and reading grade level for 7093 recruits.

In addition to these regression analyses, we are currently using the data base provided by the ability tests to evaluate various hypotheses concerning the relationship of reading ability and nonverbal intelligence. The hypotheses derive primarily from the proposals by Sticht (1972, 1975a), Jensen (1972) and others that reading ability is built upon nonverbal intellectual ability and that asymptotic reading performance is determined by this basic ability. We have the Pattern Matching test as a measure of nonverbal intelligence. In addition to this measure we will be examining factor loadings obtained in a principle component factor analysis of all our test data. As a part of this effort we will examine the contributions of verbal skill and intellectual ability to performance on the individual tests in the Basic Battery.

Reading Ability and Navy Effectiveness

A basic assumption underlying the interest in literacy in the services and the very occurrence of this conference is that reading ability is a significant factor in determining the effectiveness of a serviceman's job performance. Job performance refers to the performance of those duties associated with a position at any phase, including training, in a man's career. Logical arguments relating reading ability to job performance have been presented by Duffy et al. (1975) and Sticht (1975b). However, the conceptual and measurement difficulties involved have hindered attempts to evaluate the assumption. Conceptually, reading is viewed as an essential skill. It is one of the broad enabling skills which provide the means for acquiring specific knowledge and ability. Reading skill will play a role in determining the acquisition of specific knowledge on the job or in training as well as the specific abilities brought into the service. Because reading is so intimately related to the acquisition of specific knowledge, it is difficult to define the point at which a man's success or failure at a task is due to his specific achievement level in, for example, electronics, rather than his reading skill level. While reading is integral to the development of more specific skills, the development of reading, in turn, requires a general intellectual ability. Conceptually, general ability sets the upper limit for the development of reading skills (Sticht, 1975b). However, the independent measurement of reading and general ability has proven immensely difficult, as witness the still vexing problem of obtaining a truly nonverbal measure of intelligence (Singer, 1973).

Not only are the interrelationships of reading with other ability variables complex, but the eventual analysis of these relationships is further hindered by our inability to directly manipulate an individual's reading ability. That is, there is no program available to experimentally raise or lower an individual's reading skill by four or five grade levels, so that the effect of a change in reading can be evaluated. Rather, we must rely on correlational analyses where the performance of good and poor readers is compared. Unfortunately, the poor reader will also tend to be the individual with both low general ability and few specific skills. Thus the direct causal effect of reading on job performance is extremely difficult to extract.

Another measurement problem lies in the nature of the indices of job performance. Performance measures frequently tend to be restricted in range to scores around that point defined as "adequate." Because of this restriction in range, the magnitude of the correlation of a variable with such a performance measure will be reduced. Additionally, the considerations involved in judging performance vary from person to person and situation to situation resulting in questionable reliability. In illustration of this difficulty, Stitch, Caylor, Kern, and Fox (1971) found correlations ranging from .13 to .24 between a man's actual ability to do a job and his supervisor's rating of his capability. Since the supervisor's rating is the typical on-the-job performance measure, there is basically no reason to expect any greater correlation between reading and such a measure. Recruit training success is another job performance measure in which individual judgment and a variety of considerations are involved. However, school training performance, based on test scores, may be expected to be more reliable.

The preceding, rather pessimistic presentation of the difficulties encountered in attempting to define the role of reading in job performance is meant to emphasize the fact that correlational data alone cannot provide the necessary information. Most correlational evidence in this area can be viewed as supporting or rejecting the role of reading, depending on which of the many complex, mitigating factors are emphasized. The correlational analyses can be used to determine whether a reading test would be a useful addition to a selection battery or a job/school classification battery. In this case the interest is in performance on a reading test as a predictor of job performance regardless of the direct causal role, if any, of reading ability. However, the correlational analyses only serve as one data source for inferring the importance of reading in one job area relative to another job area. For example, if reading test performance correlated .50 with electronics technician performance but only .20 with machinist mate performance, this would suggest that reading is relatively more important in the former area. However, it is still necessary to examine the reading requirements and skill levels represented in the two areas. We need to know if large amounts of reading are expected of the men and if there is a gap between the difficulty of the material and the ability of some reasonable portion of the men. This exercise must be performed for all job areas since a low correlation of reading and performance may simply reflect the unreliability of the job performance measure in that area. Once all data sources are evaluated, inferences may be made as to the relative role of reading in a job area. However, the causal effect of reading can only be firmly demonstrated by evaluating the effects of simplifying reading material or training the reading skills of personnel.

Previous efforts to examine the reading and job performance relationship have been limited to personnel with very low reading abilities (Fisher, 1971; Hagen & Thorndike, 1953; Hoiberg, Hysham & Berry, 1974; Standlee, 1954). Inferences regarding the role of reading in these studies is tenuous at best since the "good" readers are still significantly deficient in reading relative to potential job reading requirements. To adequately assess the importance of reading, some portion of the sample must exceed a skill threshold which approximates the level required on the job. In addition to this difficulty, some of the studies inferred, rather than measured, the reading skills of the "good" readers. For example, Hoiberg et al. (1974) compared Navy recruits who had been assigned to remedial reading (poor readers) to a group with equivalent general ability, but not assigned to the program (good readers). Reading scores were only available for the poor readers. The comparison group was assumed to have adequate reading skills because they were not assigned to remedial reading. However, at the time of the study the reading skills of an individual were only assessed if he failed the first test in recruit training. Then, even if a man had a reading deficit he was only assigned to remedial reading if space permitted. Given these conditions and the equation of the groups on general ability, there is simply no assurance that the group differed in reading skill.

Our recruit and "A" School samples provide the opportunity to examine the relationship of performance on a reading test and job performance where the full range of reading skills is directly assessed and represented. Our performance measures are "A" School performance scores and attrition from recruit training. Prediction of these performance measures has been the subject of considerable research since they are fundamental considerations in the selection and classification of personnel. Thus our data, in addition to providing basic information on the role of reading in the Navy, are relevant to considerations of the use of a reading test in the selection and classification of personnel. The focus in the analyses is on the contribution of reading test performance to job performance relative to those measures already in use.

Method

Recruit Sample.

The sample of recruits for this analysis consisted of 20,627 men entering RTC San Diego between 1 June 1974 and 1 February 1975. The dependent variable was attrition from recruit training. The attrition lists at RTC San Diego were checked monthly and all attritees for whom we had reading data were noted along with the reasons for attrition. The predictor variables were:

1. Odds for Effectiveness (OFE). An estimate of the probability of completing the first tour in the service and being recommended for reenlistment. This is an actuarial table used by recruiters in which a composite score is derived from years of schooling, expulsions and suspensions from school, and Armed Forces Qualification Test Score (Plag, 1968).

2. Armed Forces Qualification Test (AFQT). A score derived from performance on the GCT, ARI, and MECH tests taken at the recruit station. Placement into a mental category is based on this score.

3. Reading Ability. Performance on the Gates-MacGinitie reading test.

"A" School Sample. The "A" School sample and the predictor variables are as described in the previous section. The dependent measure for this portion of the project was average performance on a weekly paper and pencil test. Subsequent analyses will examine an "in school" job performance measure as well as the final score in the school, which is a score based on all previous testing.

Results

Attrition. In our total sample of recruits, there was a 7.7 percent attrition rate, compared to an actual attrition rate at San Diego of approximately 10 percent during our test period. This disparity reflects the attrition of personnel prior to administration of the reading test as well as personnel who were hospitalized at the time of testing and eventually attrited. The distribution of reading ability for the attritees and nonattritees in our total sample is shown in Figure 5. The primary reasons for attrition in our sample were inaptitude (academic and military), physical disability, and psychological disability. Approximately 30 percent of the attritees were in each of these three categories. Even with this wide variability in the reason for attrition, the data in Figure 5 indicate very different distributions of reading skills for attritees and nonattritees (median reading levels are 8.2 and 10.9 respectively). The probability of attrition in each of the reading categories is shown in Table 2. A clear and systematic relationship between reading ability and the probability of attrition is evident. The less than 4th grade readers have a .36 probability of making it through boot camp, while the probability is .96 that the above 10th graders will make it.

While reading test performance is clearly related to attrition, the question still remains as to the contribution it makes relative to other available indices. It may be the case, for example, that the reading-attrition relation is dependent upon covariations with general ability. If so, reading and AFQT scores should be equally effective in predicting attrition. To answer this question, we entered the reading, OFE, AFQT, and years of education scores into a stepwise multiple regression analysis with attrition, a dichotomous variable, as the dependent measure. The sample of recruits was divided into subsamples of 15,500 and 6300 for purposes of cross-validation. The multiple regression analysis of the initial sample yielded reading as the strongest predictor of attrition, $r = .25$. The addition of the remaining three predictor variables added only .02 to the multiple correlation. The reading-attrition relationship increased to $r = .33$ in the cross-validation sample, and here the remaining three variables added only .003 to R. The independent correlations of OFE, AFQT, and years of education with attrition in the initial sample were .19, .17, and .14, respectively. The comparable correlations in the cross-validation sample were .16, .18, and .12.

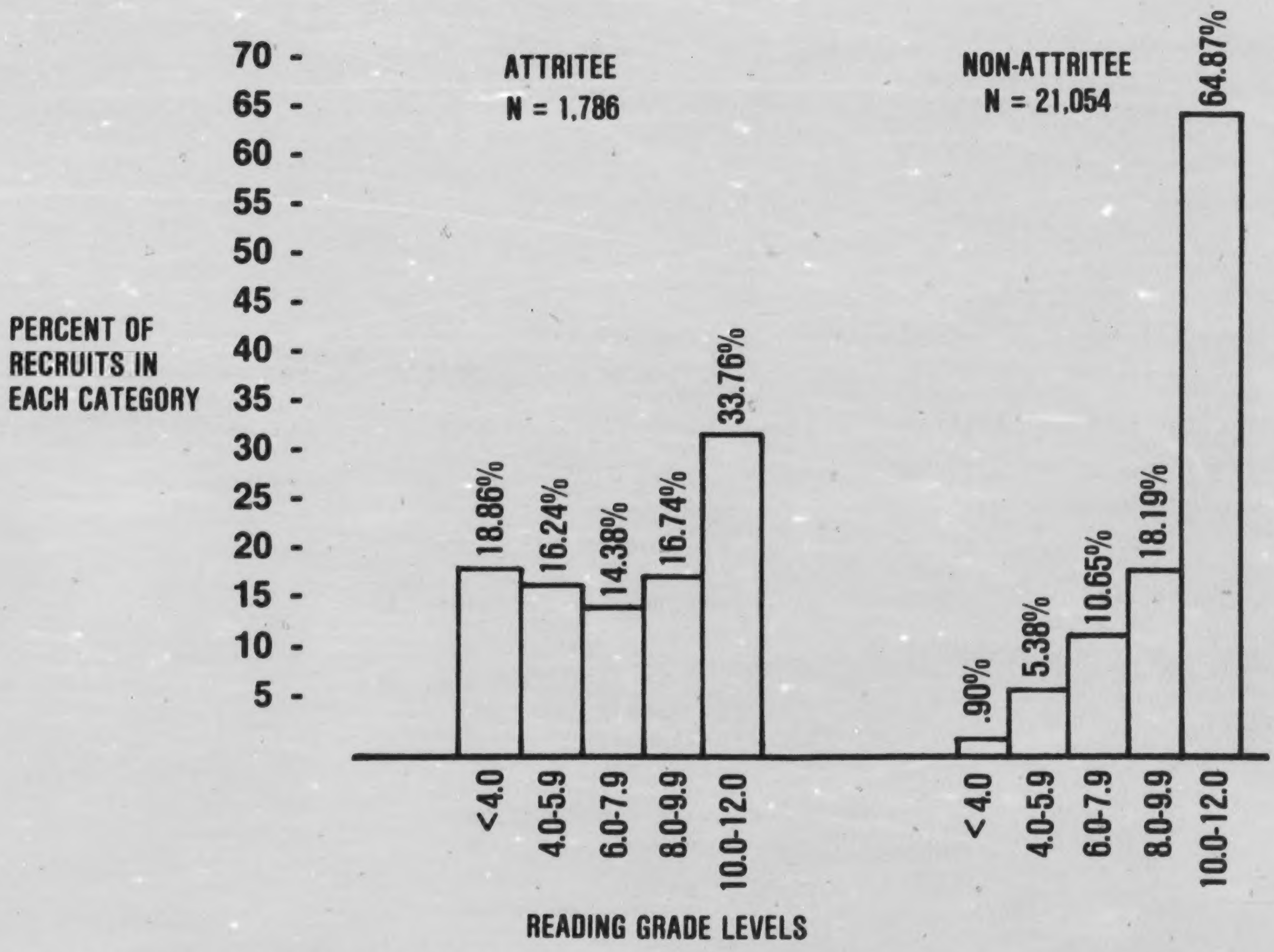


Figure 5. Reading ability for recruits attriting and graduating from RTC, San Diego, June 1974 to Jan 1975.

Table 2

San Diego Attrition Rate Within Reading Categories:
June 1974 - January 1975

	Reading Grade Levels					
	<4	4.0-5.9	6.0-7.9	8.0-9.9	10.0-12.0	Total
Number of Men	526	1424	2500	4129	14,261	22,840
Number Attriting	337	290	257	299	603	1,786
Percent Attriting	64.1%	20.4%	10.3%	7.2%	4.2%	7.81%

The attrition analyses thus far indicate that attrition rate systematically decreases as reading test performance increases, and this holds for the full range of reading levels. Additionally, of the variables examined, reading is the single best predictor of attrition within the current recruit population. It may be, however, that the contribution of reading in the prediction of attrition may be almost totally accounted for by the covariation in the AFQT, OFE, and education variables. To determine the independent contribution of reading performance to the prediction of attrition, i.e., the contribution of reading over and above the other independent variables, a second step-wise regression analysis was performed. This time, however, OFE, AFQT, and years of education were entered into the analysis before reading. These three variables together yielded a multiple correlation with attrition of .20. The addition of reading as a predictor raised this correlation to .30. Thus, reading performance contributes meaningfully to the prediction of attrition independently of the other predictor variables.

That attrition was more strongly correlated with reading performance than with the other independent variables should not be interpreted to indicate that reading performance is a more valid or important indicator of attrition. The recruits we tested had already been selected into the service on the basis of their OFE, AFQT, and education scores. Thus there is greater attenuation in the range of these scores than in reading scores and a smaller correlation is to be expected. However, reading performance clearly adds to the prediction of attrition and, therefore, the addition of a reading test to the selection battery would serve to reduce attrition rates.

The data further indicate that the reading ability of an individual is to some extent an essential skill in boot camp. That is, the fact that reading test performance contributed to the prediction of attrition independently of AFQT suggests that it is reading skill per se and not a general ability factor underlying reading test performance which is related to attrition. The data in Table 2 indicate that a reading ability of approximately the 7th grade level is necessary for achieving success. This relatively low level is consistent with the reading requirements in boot camp. That is, recruit training is primarily a process of acculturation. While reading of rather difficult material (11th grade level on the average) is required, this consumes only a portion of recruit training and all of the reading material is fully discussed in a classroom. Thus it is reasonable to assume that only basic reading skills are required to attend to routine forms, etc., and to review material presented in class.

"A" School Performance. Only a preliminary analysis of the relationship of each of our predictor variables to weekly test performance has been completed. In this analysis we were again interested not only in the degree of relationship of reading and performance, but also in how that relationship compared to the largest correlation between one of the current predictor tests and performance. The data relevant to these considerations is presented in Table 3--reading-performance correlations range from .20 to .50 across the six schools, while the maximum predictor test correlation with performance ranges from .21 to .61.

Table 3

Correlation of Reading, Pattern Matching, and the Most Predictive Test in the Basic Test Battery (BTB) With Weekly Test Average in Six Navy "A" Schools

	Weekly Test Average Correlation With:		
School	Reading	Pattern Matching	BTB Test
Hull Maintenance Technician[a]	.20	.22	.21 (ARI)
Ship's Serviceman	.26	.16	.21 (ARI)
Mess Specialist	.43	.22	.32 (ARI)
Quartermaster	.50	.48	.61 (EST)
Interior Communications	.39	.39	.50 (ARI)
Electrician's Mate	.23	.16	.44 (ETST)

[a] A weekly test is not administered in this school and therefore the daily test average score was used.

We assume that the variability across schools reflects not only variability in scoring criteria but also variability in the reading levels represented, relative amount of reading required, and reading difficulty of the material. In an attempt to find some pattern in these relationships we compared the correlational pattern in Table 3 to the pattern of reading levels and reading difficulty found in the schools (Table 1). This comparison indicates that the reading difficulty of the materials bears no relationship to either the absolute or relative effectiveness of reading as a predictor of school performance. However, reading is found to be the best performance predictor in the same three schools which have the largest proportion of low ability readers (Ship's Serviceman, Hull Technician, and Mess Specialist). A comparison of the variability of reading scores indicates that the schools are comparable in this respect (standard deviations range from 1.9 to 2.1) and thus the relationships obtained are not due to a statistical anomaly.

The finding that reading is a more effective predictor of performance, relative to other predictor tests, in schools having a large proportion of low ability readers is not surprising. The results do suggest, however, that individual differences in reading levels are most important in the lower ranges of reading ability. Similarly, a reading ability-reading difficulty mismatch is less important at the higher levels of reading and readability. More detailed data analyses and inclusion of the remaining five "A" Schools are needed to verify these clearly post hoc findings. These analyses will include multiple regressions in which reading and the current predictor of success in a school (some combination of Basic Battery tests) will be evaluated as predictors of performance. If the tentative conclusions we have drawn are verified, they suggest that lower ability readers (the specification of "lower ability" to be defined in future analyses) are poor risks in the "A" Schools. Therefore, a reading test would serve as a useful classification instrument or, alternatively, an "A" School preparatory reading program as discussed previously could be instituted to provide the deficient readers with the necessary skills.

Literacy Training

I would like now to narrow the focus of consideration to personnel demonstrating major deficits in reading skills--the below 6th grade reader. As discussed in the introduction, the Naval service provides reading training for low literate personnel at all the boot camps. I will first discuss three of these programs and the evaluation data available. In addition, I will describe three experimental training programs which have just been initiated. Finally, I will describe my own research efforts to characterize the reading deficiencies and to predict the probability of success in the current training programs.

Current Programs[6]

The three programs to be described are Navy programs in San Diego, California and Orlando, Florida, and a Marine Corps program in San Diego. The Orlando program was an experimental effort to test the Individualized Learning for Adults package developed by Research for Better Schools (RBS Program). This is a self-paced program for illiterate adults and provides training material to the 9th grade level. Training was on general literacy and focused on the specific skills of recognition of sound symbol relations (phonics), word attack (vocabulary), comprehension (literal and interpretive), and study methods. There are 129 performance objectives in the program, with 32 of the objectives devoted to decoding skills, 21 to word attack, and 41 to comprehension skills. Five entrance tests were used to determine the appropriate starting level for an individual and performance during training was regularly assessed using 70 different pre- and posttests. The training materials were 129 study booklets and 54 audio cassettes especially prepared for this program. The maximum student-instructor ratio was 7 to 1 (4 instructors and a maximum of 28 students in a classroom). The instructors were Navy men selected on the basis of having a college degree (area of study unspecified). Instructional time ranged from 10 to 160 hours with a mean of 60 hours.

---

[6] Program descriptions and evaluation data were obtained from the following sources: Marine Corps (Stewart, 1974; 1975); Navy in San Diego (Academic Remedial Training, 1975); Navy in Orlando (Research for Better Schools, 1974).

A total of 43 male recruits participated in the RBS program between September 1973 and February 1974. The criterion for selection was a score of 7.0 RGL on the combined vocabulary and comprehension subtests of the Gates-MacGinitie Reading Survey D. The methods of selecting recruits to test was not specified. After three weeks in the program, the student was retested on an alternate form of the Gates-MacGinitie test. If this score exceeded a 7.0 RGL the man was returned to recruit training, otherwise he continued through the entire RBS program.[7]

The Marine Corps (MC) program in San Diego was devised especially for the Marines in conjunction with the local adult school. This is the only program involving civilian instructors, all of whom are female and hold teaching credentials and adult school certificates. The program involves some self-pacing but is classroom based and fixed at four weeks (120 hours) of instruction. The training is in general literacy, and instructional time, according to the program syllabus, is distributed as follows: decoding (phonics), 40 percent; reading speed and comprehension, 42 percent; vocabulary development (sight vocabulary and word attack), 15 percent. Thus there appears to be a greater emphasis on decoding skills in this program than in the RBS program where 25 percent of the objectives deal with decoding. A wide variety of commercially available training materials and diagnostic tests are used in the program, including the Controlled Reading Skill Development Series (Educational Development Lab), McCall-Crabbs series (Columbia University), Reading Attainment System (Grolier Educational Corp.), Reading for Understanding (Science Research Associates), and Phonic Word Blend Flip Charts (Kenworthy Educational Service).

All recruits at the Marine Corps Recruit Depot in San Diego who read below the 4.5 grade level are accepted into the MC program. Initial screening is accomplished through the administration of the Gates-MacGinitie Reading Survey D to all recruits scoring below approximately the 70th percentile on the general classification test. Recruits scoring below the 4.5 RGL are retested on an alternate form of the Gates-MacGinitie and only if they once again score less than a 4.5 RGL are they admitted to the program. The score on the second test serves as the person's entrance score. Evaluation data are available on 492 recruits entering the program between February 1974 and July 1975.

The Navy program in San Diego (NSD program) is the only program of the three under discussion which is run entirely by the service. Instructors are Navy personnel having a college degree and the training materials are selected or developed by these instructors. Approximately 25 percent to 30 percent is devoted to phonics training. The phonics materials are based on the Motts Phonic System (unreferenced) and are similar to those used in the MC program except that phonics is concentrated in the first week of training. Successive weeks of the program deal with vocabulary (60 words per week), comprehension, and reading speed. The program is lock step and is a minimum of 3 weeks. Failure in any week results in repeating that week. Thus, the program averages 4 weeks (92 hours) and ranges from 3 to 6 weeks. Beginning in July of 1974 (FY 1975) specific testing for knowledge of sound-symbol relations was instituted. If skills in this area were adequate, then the recruit bypassed the first week. Thus, in FY 1975 the average program duration was reduced to somewhat less than three weeks (69 hours) with a range of two to five weeks.

---

[7] The fact that a student was not retested until 3 weeks into the program appears inconsistent with the report that some men completed the program in 10 hours. No explanation is available.

The training materials are a mixture of instructor-generated worksheets and commercially available supplemental materials. Hardware is limited to reading pacing devices. Prior to FY 1975 the supplemental materials consisted of the McCall-Crabbs readers and a few magazines (e.g., Scholastic Magazine). In FY 1975 a wide variety of commercially available training packages, books, and magazines were introduced into the program. It is estimated that in both FYs 1974 and 1975 approx--imately 70 percent of the vocabulary training material and more than 90 percent of the comprehension training material focused on general literacy.[8]

The NSD program is directed at recruits with an RGL between 3.0 and 5.5. Prior to FY 1975 all recruits failing the first academic test in recruit training were administered the Gates-MacGinitie Reading Survey D. If a recruit scored below the 3.0 level he was recommended for discharge. Recruits between a 3.0 and 5.5 RGL were admitted to the NSD program. In FY 1975 the Gates-MacGinitie reading test was administered to all recruits. Those scoring less than 5.5 RGL were then retested on an alternate form. Only those scoring between 3.0 and 5.5 RGL on the retest were admitted into the program.

The distinction between FY 1975 and FY 1974 procedures is important in the consideration of the effectiveness of the program. The RBS program and the NSD program prior to FY 1975 admitted students on the basis of a single test administration. Since only low scoring individuals were admitted into the program and later retested, the change scores should be subject to considerable upward regression. The MC program and the FY 1975 NSD program, by selecting students after two test failures, should reduce this regression effect. Thus, the regression effect should artifactually increase the amount of gain in the FY 1974 NSD and RBS programs to a greater degree than in the other two programs.

---

[8] The NSD program is currently undergoing another revision. The new materials are estimated to be 70 percent to 80 percent Navy and the instruction will be individually paced.

In addition to the program characteristics described thus far, differences in student characteristics and program policy may be expected to affect gains in reading within the programs. Tables 4 and 5 summarize the educational and racial characteristics of recruits entering each of the programs. The racial distribution is about equivalent in the MC and NSD programs except for an increase in the proportion of Filipinos (and thus second language training requirements) in the FY 1975 NSD program. Educational differences, however, are considerable. Approximately 50 percent of the NSD students have a high school diploma while only 20 percent of the students in the other programs are graduates. If a higher level of educational achievement is indicative of better developed study skills, then the NSD programs may show greater reading gains simply due to the relatively greater level of educational achievement of their students. Parenthetically, it is interesting to note that the data in Table 5 reflect the conclusion drawn earlier that a high school diploma does not insure that a person is functionally literate.

Table 4

Percent of Recruits in Each Racial Category for Four Reading Training Programs

	Race			
Program	Caucasian	Black	Malaysians	Others
RBS[a]				
MC	45	29	16	20
FY74 NSD	44	26	12	18
FY75 NSD	40	20	26	14

[a] Race data unavailable for this school.

Table 5

Percent of Recruits at Each Level of Education For Four Reading Training Programs

	Education Level				
Program	8	9	10	11	12+
RBS	7	16	26	30	21
MC	12	22	26	26	19
FY74 NSD	2	12	16	25	45
FY75 NSD	3	6	15	23	53

Table 6, while summarizing instructional gains in the programs, also contains information relevant to our consideration of the relative gains one might predict for the programs. First, the entry level reading scores at MC are considerably below those in the other programs. The values were not subjected to statistical comparison because of the lack of adequate raw data and because of the widely varying "n's." However, the entry scores fall within that range of ability at which phonic skills are acquired. Thus a 1 RGL difference in entry levels should reflect a meaningful difference in phonic skills, which are considered by instructors as the most difficult and time-consuming skills to acquire. In a post hoc vein, our own data and data from the MC and NSD programs indicate that entry reading level is a good predictor of amount of gain (a positive relationship). Therefore, on the basis of entry level scores the MC program should yield smaller gains. Table 7 also indicates that the NSD programs attrite a considerably greater proportion of students. Since only graduates are entered into the posttest calculation, this greater attrition should artifactually increase the gain in reading skills found in the NSD programs.

Table 7 summarizes those program characteristics which might be expected to benefit the amount of gain found in each program. The RBS program may be characterized as "program" oriented since it involved carefully developed training materials and procedures. The MC program, in contrast, capitalizes on instructor capabilities and extended contact with the instructors (a relatively long training period). Finally, the NSD program benefits are student based in that better students (higher RGL and level of education) enter the program and there is a more liberal policy of attriting students not showing progress or students demonstrating improper attitudes.

Table 6

Mean Reading Grade Level on Pre- and Posttests,
Gain in Reading, and Attrition Rate for Four Reading Programs

	Reading Scores			Attrition
	Pretest	Posttest	Gain	
RBS[a] (N=43)	4.5	6.3	1.8	.00
MC (N=490)	3.5	5.7	2.2	.10
FY74 NSD (N=785)	4.2	6.0	1.8	.22
FY75 NSD (N=658)	4.2	6.1	1.9	.24

Note. Reading was assessed on the vocabulary and comprehension subtests of the Gates-MacGinitie Reading Survey D.

[a] The "N" does not reflect attrition from the programs.

Table 7

Program Characteristics Which Are Expected to Result In Relatively Greater Reported Gains in Reading

	Program			
Characteristics	RBS	MC	FY74 NSD	FY75 NSD
Greater Instructor Qualifications		X		
Longer Program Duration		X	X	
Greater Training Materials Expenditure	X			
Greater Individualization	X			
Higher Entry Reading Levels	X		X	X
Higher Education Levels			X	X
Higher Rate of Attrition From Pre- to Posttest			X	X
Greater Statistical Regression From Pre- to Posttest	X		X	

The gain scores as well as the posttest scores presented in Table 7 indicate that despite considerable differences in program orientation, all of the programs yield approximately the same amount of gain and terminal level of reading performance. Potentially, training of some of the men in the RBS program was terminated after 3 weeks due to achieving a 7.1 RGL or higher on a test. However, all other men as well as the men in the other 3 programs were required to complete a fixed course of instruction. Additionally, in all programs, except perhaps the FY 1974 NSD program, training materials went well beyond the 6th grade level. Thus the similarity between programs in the final level of achievement does not seem to be due to a ceiling effect created by a commonality in training criteria. The difference between the programs in terms of amount of gain is only .4 RGL. Considering the substantial differences in characteristics between the programs, this small a difference in effectiveness suggests that within the confines of a short duration, instructor based, low literacy program, the amount of instructional gain in general literacy is fixed and determined by the simple exposure to the educational setting. This conclusion is further supported by the 1.2 RGL gain in general literacy obtained in the Army's Functional Literacy Training program (Sticht, Caylor, Fox, Hawke, James, Snyder, & Kern, 1973). This program, which underwent considerable experimental development, employs qualified teachers, is highly individualized, and focuses on functional reading skills. Nonetheless, it possesses the defining characteristics for my conclusion and the gain in general literacy is comparable to the gains obtained in the Naval service.

Results obtained by Shennum, Aiken, and Thomas (1975) suggest that this invariance in reading gains may even apply to specific instructional procedures. These investigators examined three procedures for increasing reading speed for recruits in the final week of the FY 1975 NSD program with the hypothesis that excessively slow reading speeds (frequently found in these programs) may be an important contributor to poor comprehension. The training method of primary interest involved simultaneous reading and listening with the rate controlled through a variable rate speech compressor. Comparison conditions were: (1) simply listening to the compressed speech; (2) reading with no listening but with rate goals specified and feedback given. Recruits in these conditions spent 2 hours per day for six days reading a 40,000 word novel written at the 8th grade level. A pretest-posttest evaluation using different materials indicated that each condition produced an increase in reading rate of about 60 words per minute and a 20 percent increase in comprehension. These gains, while constant across conditions, were nonetheless due to training since a no-treatment control showed no gain in rate or comprehension over the 6-day period.

The general conclusion I have drawn from these evaluations should not lead to the assumption that literacy training programs cannot be improved. A consideration of alternative criteria for training success (the criteria are currently the amount of gain and the exit level in general literacy), a shift to training of men with less deficiency, or the introduction of refresher or retraining programs may all yield more effective reading programs. Perhaps the most important consideration is a rethinking of the criteria for training success. In most Navy training, program effectiveness is heavily weighted by the time required to reach a criterion. The time factor is very important since students and staff are all on salary. If a time criterion were applied to the reading programs, then individualized, computer-based programs employing mastery learning techniques probably would be by far the most cost effective (Atkinson, 1972; Ball & Jamison, 1973; Bloom, 1974). The only time information available on the present programs comes from RBS, where it was found that, on the average, 41 hours of instruction were required for a one-year gain in reading. However, that one-year gain required 52 hours of instruction for recruits without decoding skills, but only 25 hours when recruits could decode.

The criterion for literacy training should also be rethought in terms of the ultimate goals of the program. The goal is not to produce more literate men but rather to give each man the skills necessary to effectively perform the reading tasks required of him in the service. These skills involve an ability to readily find information in manuals, to read and comprehend procedural directions and instructions, to determine the relative importance of information on a system, etc. The vocabulary requirements and the reading strategy required in these tasks differ considerably from the requirements in general literacy. Tests based on service-relevant reading material and tasks will very likely show program gains not reflected in general literacy tests.

Still related to the ultimate criterion of increased Navy effectiveness, a program addressing men with minor reading deficiencies may be expected to yield greater gains in performance effectiveness as compared to low-literate programs even if both programs produced the same gain in the same time. While the advantages of the former program were discussed earlier, no information was available on the relationship of reading improvement to entry-reading level. Data derived from these literacy programs indicate that, in addition to the gains in Navy effectiveness, a program for moderately literate men would require less training time to achieve the same gains in reading. The NSD personnel refuse to accept less than 3.0 RGL readers as they have found training of these men to be virtually impossible. RBS reports that when a man is not proficient in phonic skills he requires twice as much training to produce a year gain in reading. In my own work at NSD I have found a strong linear relationship between entry level and gain ($r = .35$ and $.59$ in two samples of native English speakers).

Literacy training programs have frequently been criticized for the tendency to be one-shot programs. Clearly meaningful permanent gains in reading cannot be expected in a three-week program. Exercising of the new skills and refresher training are necessary. The NSD personnel retested 58 graduates from the FY 1975 program 2 to 7 weeks after graduation. The delay between the final test and retest did not produce any systematic differences in change scores, which indicates that the loss due to termination of training occurs within two weeks. The average graduation score for these 58 men was 6.3 RGL while the average retest score was 5.1, a 1.2 RGL loss in reading skill due to the termination of training. If these men were representative of the FY 1975 input into the program, these data indicate a .9 RGL permanent gain from the entry reading level of 4.2 RGL. However, this gain is still subject so some regression effect and so the permanent gain in the program is likely a .5 to .8 RGL. While retest data is not available on the other Navy programs there is no reason to suspect the permanent gain would be any greater than that obtained at NSD. The exit reading score indicates that the men have the capability of reading above the level attained in retesting. To achieve that higher level on a permanent basis, however, will require mini-refresher courses over an extended time frame.

That the reading programs produce approximately equivalent gains and that there is considerable loss in skill after a man leaves the program should not detract from their relative effectiveness. The pretest and posttest gain of approximately 2 years in these compulsory and time-compressed programs is indeed significant. Looking at the gains in another way, the men, on the average, ranked at the 30th percentile for 5th grade students at entry into training, while on the posttest, the men were at the 69th percentile. Thus, during the course of approximately four weeks of training, these men move from the lower to the upper third of the distribution for this particular reference group. The 2 RGL gain is also significant relative to an average .5 RGL gain achieved in civilian adult school programs of the same total instructional hours but spread over five months (and therefore providing distributed practice) and offered on a voluntary basis (Kent, 1973). Similarly, experimental reading programs for school children at the same reading skill level as the recruits have produced less than a year gain in a year of instruction (Battelle Institute, 1972). In comparison, the Naval service programs have proved exceedingly effective in improving reading skills. My discussion, instead, addresses whether more effective programs can be developed and whether the programs produce gains to the service beyond the increase in reading.

With regard to the effects of the reading programs on later Navy effectiveness, graduates of the NSD and MC programs have been tracked through boot camp to determine their rate of attrition. The success of a sample of 301 MC program graduates was assessed 3 to 14 months after graduation. The attrition rate of this sample was 12 percent, compared to an overall boot camp attrition rate of 10 percent. Since boot camp is only 11 weeks, these data indicate that graduates of the MC program have an attrition rate equivalent to or less than the Marine Corps average. Additionally, only 33 percent of the graduates were judged to have a promotion rate slower than normal, while for 17 percent the rate was judged to be faster than normal. Unfortunately, an untrained sample is not available for comparison purposes. (Perhaps most men are typically judged to be progressing faster than normal.) The data do suggest that the MC program graduates are progressing through the service at a normal rate. However, we do not know to what degree reading ability is predictive of performance in the Marine Corps (it is evident that the reading requirements are less than in the Navy). Thus the "normal" performance of these graduates may be due to some characteristic of the reading program or it may be due to the lack of any relationship between reading ability and service effectiveness.

Boot camp attrition rate for a sample of 387 graduates from the FY 1974 NSD program was 12 percent. This can be compared to an average 10 percent attrition rate at the boot camp and an attrition rate of 18.7 percent for 4th and 6th grade readers. As with the MC program, the graduates of the NSD program appear to perform as well as the average recruit. Here, however, we have data pointing to a relatively strong relationship between reading and attrition (see Table 2) and thus the 12 percent attrition rate for graduates appears to be significant.

The graduates of both the MC and NSD reading programs appear to be performing as average recruits. However, a basic question is whether their level of performance in the service is due to the improved skill level and attitude obtained from the reading program. While a definitive answer to this question is not possible, it seems highly unlikely that a 1.0 RGL increase in ability would substantially affect the performance level of a functionally illiterate recruit. A more likely effect of the programs is that they serve to filter out recruits having low capability for learning or having a poor attitude. These people are represented in the 12 percent and 14 percent attrition rates in the MC and FY 1974 NSD programs, respectively. The only way we are going to be able to adequately assess the effectiveness of reading training on later service effectiveness is through an adequate experimental design in which only a portion (randomly selected) of the personnel eligible for reading training are actually assigned to the program and the program resulted in meaningful gains in reading skill. A judgment of reading training effectiveness would then be based on attrition rates from the point of assignment, rather than examining only the graduates of the reading program.

Experimental Reading Programs

In addition to the ongoing programs for recruits there are three experimental programs under Navy sponsorship. Since these programs are only in the initial stages of development, I will only briefly describe them. Researchers at NAVPERSRANDCEN are developing programs for training phonic or decoding and vocabulary skills in the low-literate recruit. The training is computer based and has as dual goals the evaluation of training procedures and the evaluation of new computer hardware and software systems. The interactive computer system has graphic capabilities, allowing the student to "point" to stimuli, e.g., to syllabicate a word shown on the screen. More importantly for the low-literate training, the terminals are equipped with a Vortax voice synthesizer. This computer-generated speech capability permits a fully interactive and individualized training program with computer presentation in either the auditory or visual mode or both. A programming language incorporating the graphic and voice synthesizer capabilities has recently been generated and the courseware for the phonic skill training is nearing completion.

The second experimental program is being developed by Carver (1973a) under the sponsorship of the Office of Naval Research. The training in this program, rather than focusing on specific reading skills, will involve 150 hours of practice on the terminal objective--reading and comprehending prose material. Prose passages will be presented over a PLATO computer terminal, thus permitting individualized instruction. The passages will be presented in a modified cloze format (Taylor, 1957) in which every fifth word is deleted and the student makes a multiple choice selection to fill the blank. This procedure forces continued attention to the passage and provides constant monitoring of performance. The primary objective in Carver's work is to determine if there are two types of poor readers--those lacking the intellectual capacity and those lacking the educational experience. The latter are expected to gain from reading training. Thus, students will be screened for intellectual ability using the Raven's Progressive Matrices and the high and low scorers will be compared. Those receiving training will consist of 4th grade level readers. Piloting testing of this system has recently begun.

The final experimental program involves the evaluation of the Encyclopedia Brittanica program for functionally illiterate adults. The program is at a considerably lower level than the other programs we have discussed and assumes no reading ability. Training is in pronunciation, listening, and basic reading skills. This program has been instituted aboard three ships where participation is voluntary. The experimental evaluation of this program, by the Chief of Naval Education and Training Support, is just getting underway.

Prediction of Training Success

As a final topic in the review of the Navy's literacy research, I would like to present some of the results obtained in my own research efforts to characterize the successful student in the NSD reading program. For the past several months we have been administering a battery of tests to all recruits entering the program and examining test scores in relation to the amount of gain in reading ability.[9] Our first sample consisted of 111 recruits, 32 of whom had English as a second language. We then made some modifications in the test battery and have recently begun collecting data on a second sample. This sample, at present consists of 41 recruits, 21 of whom have English as a second language. It should be borne in mind that the data I am presenting applies to the 1975 NSD program which was described previously. Programs differing in instructional strategies would likely produce different relevant variables.

One of the first hypotheses we examined was that the remediability of poor readers is dependent upon their intellectual ability. This is a basic hypothesis in Carver's (1973) experimental reading program discussed previously. To test this hypothesis we used the pattern matching (PM) test (derived from the Raven's Standard Progressive Matrices as described previously) as a part of our battery. The split half reliability (corrected for attenuation) of the PM, based on the scores of 1200 recruits, is .85. Our dependent variable was the gain in reading scores over the course of the reading program. If a student was attrited from the program he was given a score of 2.6, .4 RGLs below the lowest exit score. Mean PM performance was consistent across the four groups (successive samples crossed with English as a 1st or 2nd language) ranging from 19.5 to 22.5 (σ = 4.9 to 6.6). This is significantly below the mean of 28.3 obtained for all recruits at the San Diego boot camp. The mean gain in reading ranged from .8 to 1.8 RGLs but was not consistently related to a language group. The gains and the correlation of gain and PM performance is presented in Table 8 for the English as a native language (E 1st) and English as a second language (E 2nd) groups.

---

[9] Analysis involving posttest scores have yielded comparable results.

Table 8

Mean (Standard Deviation) Reading Gains[a] and the Correlation Of Gain With Predictor Variables for Native Language (E 1st) And English as a Second Language (E 2nd) Groups in Two Samples

		Correlation of Gain With:		
	Group (N)	Pattern Matching	Decoding Skill	Mean Gain
Sample 1	E 1st (79)	-.02	.40	.8 (1.7)
	E 2nd (32)	.19	.30	1.8 (1.9)
Sample 2	E 1st (21)	-.39	.54	1.23 (1.49)
	E 2nd (29)	.18	.28	1.0 (1.1)

[a] Reading gain is calculated as the posttest-pretest difference on alternative forms of the vocabulary and comprehension subtests of the Gates-MacGinitie Reading Survey D.

The correlation between PM and gain is low and not statistically significant ($p > .05$) across all four groups. If anything, there is a tendency toward a negative correlation between gain and pattern matching for the E 1st group. Thus intellectual ability of low-literate recruits is not predictive of their gain in this reading program. This result is consistent with our finding with a random sample of all recruits, where it was found that PM was not strongly related to general reading ability ($r = .43$). It seems that while intellectual ability is a component of reading skill, it alone does not set the limits for the level of skill attained in a reading program, at least as found in our heterogeneous adult population.

Our second hypothesis was that those students more proficient in decoding printed words into phonological units would show greater gains in reading. This is often reported by the instructors in the reading program and, in the NSD program, it is the rationale for rejecting recruits below the 3.0 RGL. The word knowledge subtest of the Wide Range Achievement

Test was included in our battery to test this hypothesis. The word knowledge test involves the presentation of a graded series of printed words with the subject's task being to simply pronounce the words correctly. Correctness of pronunciation is a test administrator judgment and grading is liberal with respect to regional and ethnic differences in pronunciation. The test does not take into account the student's knowledge of the meaning of the words.

Results obtained with the decoding skill test supported the hypothesis. Performance on this test was consistently one of the best predictors of gain, with the correlation being somewhat higher for the E 1st sample ($r = .40$ and $.54$) then for the E 2nd sample ($r = .28$ and $.30$). As indicated by the instructors, decoding skill is an important determinant of ability to learn to read in a short-duration program. Decoding skill, however, is clearly not the only factor. The mean decoding performance (in grade levels) for the E 2nd samples was 10.1 as compared to a mean of 4.7 for the E 1st, a difference of 2 standard deviations. The magnitude of this difference was consistent across samples, yet the gain in reading for the E 2nd group is not consistently greater than for the E 1st group. We feel that a second major distinction between the language groups, which will account for differences in gains is semantic knowledge. The E 2nd consists primarily of Filipinos who have been taught decoding rules for English since grade 1. However, their experience with the language has been largely limited to schoolroom use. Thus, while they can readily decode the printed word, their semantic knowledge is deficient relative to the E 1st and will affect their progress in the remedial program.

We included reading and listening tests in our battery to test the above hypothesis as well as to test the hypothesis that a large difference between reading and listening skills is predictive of the gains to be made in a remedial program. The tests were developed from the vocabulary subtest of the Nelson (1962) reading test. The odd numbered items on the test formed our reading test and the even test items formed the listening test. Presentation of reading and listening items was alternated in blocks of 5. "Listening" test is somewhat of a misnomer since both reading and listening items were presented via slides at a 12 sec rate. For the listening items, however, the stimulus and alternatives were read to the subject while he viewed them. Split-half reliability of the listening test calculated on our first sample yielded an $r = .79$ (corrected for attenuation).

The mean listening and reading performance for each language group is presented in Table 9. The data indicate, as might be expected, that E 1st recruits have a larger listening than reading vocabulary. Their listening skill is exercised constantly through interaction in an English language community, while their reading skill training is likely to be limited to use in the educational system. The E 2nd recruits, on the other hand, have equivalent reading and listening vocabularies indicative of an exposure to English, spoken or written, which is limited to the classroom. The data further indicated that the E 2nd recruits have a vocabulary knowledge greater than the reading vocabulary but less than the listening vocabulary of the E 1st. The E 2nd recruits, while being able to decode far more words than the E 1st group, have more limited semantic knowledge of those words.

Table 9

Mean (Standard Deviation) Reading and Listening Ability Scores for Native Language (E 1st) and English as a Second Language (E 2nd) Groups Collapsed Across Two Samples

Group	Reading	Listening
E 1st	26.1 (5.1)	31.1 (5.0)
E 2nd	29.7 (5.0)	28.0 (4.8)

We have calculated the correlations between reading and listening skills and gain scores. However, the pattern of correlations is not stable or clearly interpretable. Thus, I will delay discussing that data until our second sample is complete. We have included a number of additional measures in our battery for the second sample. These include some of the traditional measures of Navy effectiveness, such as number of dependents and years of education, as well as tests of immediate memory and ability to maintain a high level of performance at a tedious task. The intent of these tasks is to look for what may be general ability deficits in our low-reading ability recruits.

References

Academic Remedial Training San Diego: Comprehensive Document: San Diego, California: Academic Remedial Training Unit, Recruit Training Command, 1975.

Atkinson, R. C. Ingredients for a theory of instruction. American Psychologist, 1972, 27, 921-931.

Ball, J., & Jamison, D. Computer-assisted instruction for dispersed populations: System cost models. Instructional Science, 1973, 1, 469-502.

Battelle Memorial Institute. Interim report on the Office of Economic Opportunity experiment in educational performance contracting. New York, New York: Author, January 1972.

Biersner, R. J. Reading grade levels of Navy rate training manuals and non-resident career courses (CNETS Report 2-75). Pensacola, Florida: Chief of Naval Education and Training, May 1975.

Bloom, B. S. Time and learning. American Psychologist, 1974, 29, 682-688.

Carver, R. P. The relationship between intellectual capacity and reading competency. A Research Proposal. Silver Spring, Maryland: American Institutes for Research, September 1973. (a)

Carver, R. P. Measuring the reading ability levels of Navy personnel. Technical report under contract to the Psychological Science Division, Office of Naval Research, October 1973. (b)

Carver, R. P. Measuring the reading difficulty levels of Navy training manuals. Technical report under contract to the Psychological Sciences Division, Office of Naval Research, October 1973. (c)

Caylor, J. S., Sticht, T., Fox, L., & Ford, J. Methodologies for determining requirements of military occupational specialties (HumRRO Technical Report 73-5). Alexandria, Virginia: Human Resources Research Organization, March 1973.

Duffy, T. M., Carter, J. D., Fletcher, J. D., & Aiken, E. G. Language skills: A prospectus for the naval service (Special Report SR 76-3). San Diego, California: Navy Personnel Research and Development Center, October 1975.

Fisher, A. H., Jr. Army "new standards" personnel: Effect of remedial literacy training on performance in military service (AFHRL TR 71-13). Alexandria, Virginia: Air Force Human Resources Laboratory, Manpower Development Division, April 1971.

Gunderson, E. K., & Ballard, K. B. Prediction of delinquency in naval recruits (prepared under ONR Contract NoNr 1535(00)). San Diego, California: United States Retraining Command, undated.

Gates, A. I., & MacGinitie, W. H. Gates-MacGinitie Reading Tests, Survey D. New York: Teachers College Press, 1965.

Hagen, E. P., & Thorndike, R. L. A study of World War II Navy careers of illiterates sent through literacy training. Washington, D. C.: Bureau of Naval Personnel, Classification and Survey Research Branch, April 1953.

Hoiberg, A., Hysham, C. J., & Berry, N. H. The neuropsychiatric implications of illiteracy: 20 years later (Report No. 74-20). San Diego, California: Neuropsychiatric Research Unit, February 1974.

Jensen, A. Genetics and education. New York: Harper and Row, 1972.

Kent, W. P. A longitudinal evaluation of the adult basic education program. Falls Church, Virginia: Systems Development Corporation, November 1973.

Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. Derivations of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy enlisted personnel (Res. Rep. 8-75). Millington, Tennessee: Chief of Naval Technical Training, Navy Technical Training Command, February 1975.

Madden, H. L., & Tuppes, E. C. Estimating reading ability level from the AEQ general aptitude index (PRL-TR-66-1). Lackland Air Force Base, Texas: Personnel Research Laboratory, Aerospace Medical Division, February 1966.

Nelson, M. J. The Nelson Reading Test (revised edition). Boston, Massachusetts: Houghton Mifflin Co., 1962.

Nelson, M. J., & Denny, E. C. The Nelson-Denny Reading Test (revised by J. I. Brown). Boston, Massachusetts: Houghton Mifflin Co., 1960.

Northcutt, N., Selz, N., Shelton, E., Nyer, L., Hickok, D., & Humble, M. Adult functional competency: A summary. Austin, Texas, University of Texas, March 1975.

Plag, J. A. Predicting the effectiveness of naval enlistees (Tables and summary of a report). San Diego, California: Naval Medical Neuropsychiatric Research Unit, May 1968.

Plag, J. A., & Hardacore, L. E. The validity of age, education, and GCT score as predictors of two-year attrition among naval enlistees (Report Number 64-15). San Diego, California: Navy Medical Neuropsychiatric Research Unit, June 1964.

Raven, J. C. Standard progressive matrices: Sets A, B, C, D, & E. London, England: H. K. Lewis & Co., Ltd., 1958.

Research for Better Schools. A pilot model remedial education program. Philadelphia, Pennsylvania: Author, 1974.

Shennum, W. A., Aiken, E. G., & Thomas, G. S. Uses of time-compressed speech in a reading remediation program: Some exploratory tests (Tech. Rep. TR 76-13). San Diego, California: Navy Personnel Research and Development Center, September 1975.

Singer, H. IQ is and is not related to reading. Paper presented at the Annual Meeting of the International Reading Association. Denver, Colorado, May 1973.

Spearman, C. Theory of a general factor. British Journal of Psychology, 1946, 36, 117-131.

Standlee, L. S. A follow-up and comparison of three groups of Navy enlisted men: Marginal and illiterate, marginal-but-literate, and typical recruits (Technical Bulletin 54-20). Washington, D. C.: Bureau of Naval Personnel, Training Research Branch, July 1954.

Stewart, V. T. 1st quarter report of Phase II pilot program: Recruit remedial reading program (Prepared under Marine Corps Contract Number M00243-74-C-0020). San Diego, California: Community Colleges, Midway Adult School, 1974.

Stewart, V. T. 2nd quarterly report of recruit remedial reading program, phase III (Prepared under Marine Corps Contract Number M00243-75-C-0015). San Diego, California: Community Colleges, Midway Adult School, July 1975.

Sticht, T. G. Learning by listening. In J. B. Carroll & R. O. Freedle (Eds.). Language comprehension and the acquisition of knowledge. New York, New York: John Wiley & Sons, 1972.

Sticht, T. G. The acquisition of literacy by children and adults. Paper presented at the Second Delaware Symposium on Curriculum, Instruction, and Learning: The Acquisition of Reading. University of Delaware, June 1975.

Sticht, T. G. (Ed.). Reading for working. Alexandria, Virginia: Human Resources Research Organization, 1975.

Sticht, T. G., Caylor, J., Kern, R., & Fox, L. Determination of literacy skill requirements in four military occupational specialities (Tech. Rep. 71-23). Alexandria, Virginia: Human Resources Research Organization, December 1971.

Sticht, T. G., Caylor, J. S., Fox, L. C., Hawke, R. N., James, J. H., Synder, S. S., & Kern, R. P. HumRRO's literacy research for the U. S. Army: Developing functional literacy training (Professional Paper 13-73). Alexandria, Virginia: Human Resources Research Organization, December 1973.

Taylor, W. L. "Cloze" readability scores as indices of individual differences in comprehension and aptitude. Journal of Applied Psychology, 1957, 41, 19-26.

Young, D., & Jamison, D. T. The economic benefits of schooling and reading competence (Report RB 75-19). Princeton, New Jersey: Educational Testing Service, May 1975.

READABILITY RESEARCH IN THE NAVY

Thomas E. Curran

A great deal of concern has recently been voiced with regard to the quality of instructional and maintenance manuals in the Navy. Men, both on the job and in the classroom, have tended to have problems using manuals due to the difficulty of both text and graphics, inadequacies in content, lag time in updating, etc. The thrust of this paper is with the first of these deficiencies, but its concentration differs from that often found in work on readability. Readability, per se, is not a difficult problem if one assumes as its operational definition a "readability formula score." As measured by readability formulas, almost anyone can write readably. One simply uses short words and short sentences. Given these, the "readability" of the material will be at a high ("easy") level when indexed by the most common formulas. Obviously, this is a simplistic view, particularly for technical writing. Technical writers (nearly all writers preparing Navy materials) cannot always obey the dictum to use short words. What "easy" word can be substituted for "oscilloscope," for example? But even for these writers, readability need not be a severe problem. One must merely modify the above description to say "use familiar words and short sentences" and modify the readability assessment procedure so that long familiar words do not inflate the reading grade level (RGL) of the material. A major part of the R&D effort in the field of readability should therefore be directed at identifying long words which are known to be familiar to the intended audience and effectively transmitting this information to the writer. This topic will be discussed in other contexts at a later point.

There is a second characteristic of writing that deserves special attention--comprehensibility. There seems to be some misapprehension in the literature to the effect that high readability ensures comprehensibility. Only if readability is taken to include good "style," smooth flow of ideas, avoidance of complex sentence structures, etc. is this the case. Note that this represents a rather drastic departure from the operational definition of readability suggested above, and, in fact, closely resembles the accepted definition of comprehensibility. But the position is taken here that they are clearly two different concepts. At this point in time, readability formulas cannot adequately assess characteristics such as style and sentence structure. It is true that readability and comprehensibility often go hand in hand. The nursery rhyme "Mary Had A Little Lamb" is written at a very low level of difficulty, and because of the structure of its sentences, its simple words, and its easy flowing style is also quite comprehensible. But examine the other extreme (unlikely as it would be coming from any rationale author). If the words in the rhyme were "scrambled," with the length of sentences left intact, its readability score according to formula would be identical to that of the original version. Yet its comprehensibility would be reduced enormously. In between these two extremes fall countless examples of writing which vary along each of the two dimensions. In short, comprehensibility, like readability, its a property of the written material, and when the work is done carefully the two are at least moderately correlated. But bear in mind that the former cannot be directly assessed by means of existing formulas.

To further complicate the issue, there is really no point in speaking of _either_ readability or comprehensibility unless the intended audience is taken into account. Two characteristics of the audience (other than general reading ability) are of prime importance: the background knowledge possessed and the degree of motivation involved. The adult reader faced with the scrambled version of "Mary Had A Little Lamb" might comprehend it reasonably well because of repeated exposure to it in the past, and might find it to be a stimulating (i.e., motivating) puzzle-solving experience. The same strings of words seen by a child for the first time would probably be incomprehensible, even though the individual words themselves could be decoded.

Lest one be misled by the above statements, it should be made clear that readability is NOT a trivial issue. It is a _necessary_, but not a _sufficient_ condition for high-quality manuals. All other factors being equal COMPREHENSIBILITY requires high READABILITY, and one step removed, USABILITY requires high COMPREHENSIBILITY. If a manual is not readable, the student or technician will not even attempt to use it once he has been stymied by its difficulty at the outset. He probably will avoid it from that time on, just as would anyone presented with a tome replete with long words which he cannot decode. But to be readable is not enough. If the manual is not also comprehensible, the man possibly will not be able to use it despite his need for the information that it contains. And the problem again is not even that simple. Material can be virtually unreadable according to a formula score and yet one who is vitally interested in the topic and/or needs the information regardless of the effort required to obtain it may slave his way through it despite the writer's opposition.

The interactions among readability, comprehensibility, and usability (all properties of the _material_) and reading ability, comprehension ability, motivation, and background experience/knowledge (all properties of the _user_) should provide the overall model for any R&D effort in improving Navy manuals.

In the third quarter of Fiscal Year 1975, the Navy Personnel Research and Development Center (NAVPERSRANDCEN) was tasked to investigate the readability and comprehensibility of technical manuals (TMs). An extensive survey of the literature was undertaken to determine the state-of-the-art in this area and report findings (Curran, 1975). The conclusions of this author were that the area of readability and comprehensibility has been covered in breadth but not in depth. A dilemma appears to have evolved. Formulas can be applied efficiently--automation has become the rule--but it is possible that only judgments of the material by readers can provide an index of its comprehensibility.

Prediction of Readability

This author makes the same distinction between prediction of readability and production of readable writing as put forth by Klare (1975). This section deals with the first of these problems while the following section deals with the latter and the interaction between the two.

Readability Formulas: General

The most common method for predicting the difficulty of a given piece of material is the readability "formula." A great many of these have been developed, with the majority using as variables some index of sentence difficulty and individual word difficulty. Reviews of these formulas up to 1960 can be found in Klare (1963); Klare (1974-1975) reviews those developed after 1960 and modifications to earlier ones. Reference in this paper will be made to only five such formulas: the FORCAST formula, the RIDE scale, the Fog Count, the Flesch Reading Ease (RE) formulas, and the Dale-Chall formula.

In general, the prediction of readability by formula involves "counts" of various components of written material and, using passages of known difficulty as criteria, computing a regression equation on the counted characteristics. Commonly, the criterion passages are developed using the "cloze" technique. This procedure, which is of concern throughout this paper, involves extracting verbatim passages from the material and deleting every nth word, replacing each with blanks of standard length. Normally, every 5th word is deleted, resulting in five versions of the test for each passage (deletion of words 1, 6, 11, ... n, up to 5, 10, 15, ... n). Thus, each word in the passage is deleted once across the five versions. Subjects of known reading ability are asked to fill in as many of the deleted words as they can, without having first read the intact passage. The reading level at which 50 percent of subjects can fill in approximately 40 percent of the missing words is usually taken as the reading grade level (RGL) of that passage.[1] This "scaled" RGL is then used as the criterion for regression of the counted variables.

As an example, Bormuth (1969) used this procedure for deriving a number of readability formulas. He examined, in an extensive correlational study, approximately 170 different variables, such as vocabulary, syntactic complexity, and parts of speech, and their relationship to the difficulty of written material. He first determined the correlation of each of these variables with the difficulty of 330 100-word passages as scaled with the cloze technique. He then entered the variables into a regression equation to determine which of them in combination best predicted the difficulty of the passages. The outcome of this study was a series of formulas, each designed for a different purpose, incorporating the "best" of the variables for predicting the difficulty of other writing.

---

[1] RGL is comparable to, but not equivalent to, school grade; it is a somewhat arbitrary level at which a particular grade student should be able to read with satisfactory comprehension. 40 percent correct on the cloze test has been shown to be approximately equivalent to a 75 percent score on a multiple choice test on the material.

The Flesch "Reading Ease" Formula

Probably the most widely used (and most consistently powerful) of the readability formulas is that developed by Rudolf Flesch (1948). This formula--termed the Reading Ease (RE) formula--uses as variables the number of words per sentence and the number of syllables per 100 words of text as its variables. The original RE formula is presented below:

$$\begin{aligned}\text{READING EASE} = {} & 206.835 \\ & - 1.015\ (\text{words/sentences}) \\ & - .846\ (\text{syllables/100 words})\end{aligned}$$

This formula was developed in much the same manner as described above, except that the criterion was set of standardized passages rather than passages normed for the specific purpose by the cloze technique. The formula indexes the difficulty of material on a scale from 0 (practically unreadable) to 100 (extremely easy). Using this index, the RGL can be determined from a conversion table. The RE formula was developed and validated on children and civilian adults, with reading material appropriate to these samples, and its usefulness for assessing the difficulty of military technical or instructional writing is therefore suspect. In order to overcome this problem, Kincaid, Fishburne, Rogers, and Chisson (1975) recalculated the formula using Navy enlisted personnel reading Navy job-relevant literature.[2] In the process of this recalculation the RE formula was also revised to produce RGL directly without having to take the extra step of converting "reading ease" to RGL via a conversion table. The recalculated version of the RE formula is shown below:

$$\begin{aligned}\text{GRADE LEVEL} = {} & .39\ (\text{words/sentences}) \\ & + 11.80\ (\text{syllables/word}) \\ & - 15.59\end{aligned}$$

The RE formula has the advantage of being relatively simple to compute manually and is adaptable to automation as well. Klare, Rowe, St. John, and Stolurow (1969), among others, have developed a computer program which provides the RE index based on the original Flesch formula. It remains a fairly simple task to revise the program for the recalculated version of the formula.

Application of the Reading Ease Formula

Based on the RE formula's consistent validity when compared with other formulas and its relative ease of manual computation, the Chief of Naval Education and Training Support "Readability Working Group" used the formula in an analysis of all Navy rate training manuals (RTMs) to determine their difficulty. Biersner (1975) reports on this effort. He analyzed 185 Navy RTMs, using both the original Flesch formula and the revision by Kincaid et al. (1975). (As with Biersner, this formula will be referred to herein as the Kincaid formula.) Biersner's work illustrates vividly the importance of the prediction of readability in general and to the Navy in particular. Bear in mind, throughout the discussion to follow, that RTMs are a major source of information required by Navy personnel for advancement in rate. Twenty-five samples were taken from each of these manuals, their difficulty was determined, and the average for the entire manual was computed. Considering only the Kincaid formula (which is considered more appropriate for Navy writing), the range of difficulty of the 185 RTMs was from an RGL of 8.82 (for Basic Machines) to one of 16.26 (for Disbursing Clerk 1 & C) with a median RGL of about 12.6. That is to say 50 percent of the RTMs were written at a level of difficulty (as indexed by the formula) such that a "beginning college reading level" would be required to understand them. This should be interpreted with caution, however. There is no one-to-one correspondence between the difficulty level indicated by a readability formula and the ability of a person to profit from the written material. Biersner (1975) makes this point when he says, "The relationship between RGLs (as determined by any of the reading formulas which are available) and reading comprehension or performance effectiveness is not well established, despite the importance of reading to the development of most other skills" (p. 7). It may well be that the mere presence of a Navy recruit in the naval mileau may be sufficient for him to comprehend the necessary elements of the RTM required for advancement. But this does point up the issues raised in the introduction. A formula can index the readability of material, but (at least as yet) no formula can index its comprehensibility. It is clear that the value of writing can be accurately determined only from within the framework of an overall model--the readability and the comprehensibility of the material and the state of the user in terms of motivation, background knowledge, and experience, etc. But at face value, assuming all other factors to be equal, the comparison between the RGL of the manuals as reported by Biersner and the reading abilities of recent Navy recruits (Duffy & Nugent, Note 1) shows a quite apparent "man-manual mismatch." This comparison is presented in Figures 1 and 2. Figure 1 indicates the full range of the RGLs of the RTMs from the one manual at the 8th grade RGL to those at the 14th and above. Reading abilities of recruits (as measured by the Gates-MacGinitie test) are shown for the same intervals. The fact that no reading abilities are shown for levels above 12th grade is an artifact, due simply to the fact that the ceiling for this test is at the 12.0 level. Figure 2 presents the same data in a different perspective. Here, reading abilities are indicated by percentages from grade 4.0 (and below) to grade 11.0 (and above). All RTMs whose difficulty exceeded 11.0 according to formula score are bulked into the latter interval. Noting the cumulative percentages (again at face value) it would be possible that approximately 24 percent of these recruits could read only one RTM--the single manual in the sample below the 9.0 level. To interpret further, approximately 82 percent of the RTMs are written at the 11.0 level or above; only 45 percent of the recruits were tested as reading at the 11.0 level or above. And one further illustration of the mismatch (not indicated in the figures) deals with the manual *Basic Military Requirements*. The RGL of this manual is 10.85. Knowledge of its contents are required for advancement to *E-3*. Yet its difficulty level is approximately the same as the *median* reading level of the over 21,000 recruits. It is possible that 50 percent of the entire sample would not be able to read this manual. Clearly, our ability to predict the difficulty level of written materials is of immense value. The data discussed above gives us every indication that certain RTMs should probably be revised downward in difficulty level unless reading abilities can be increased.

---

[2] These authors also recalculated two other formulas--the Automated Readability Index and the Fog Count--which will be addressed at a later point.

In addition to the striking findings reported above, another feature of Biersner's work deserves mention. As noted earlier, the computation of the reading ease formula involves a count of syllables, words, and sentences. While this can be accomplished relatively simply with nothing more than a paper and pencil, Biersner reports on a device which greatly facilitated the giant task of his analysis. Biersner directed the development of an electro-mechanical counting device which would permit more rapid and reliable gathering of the data. This device took the form of a "stylus" which, when pressed to a surface, trips a microswitch, which in turn activates a counter. A relatively unskilled person using this device can therefore press the stylus to the working copy at the end of a syllable, word, or sentence, and that variable would be indexed by the counter. Biernser (1975) reports that "It made data collection over 30 percent faster, while maintaining high reliabilities" (p. 17).

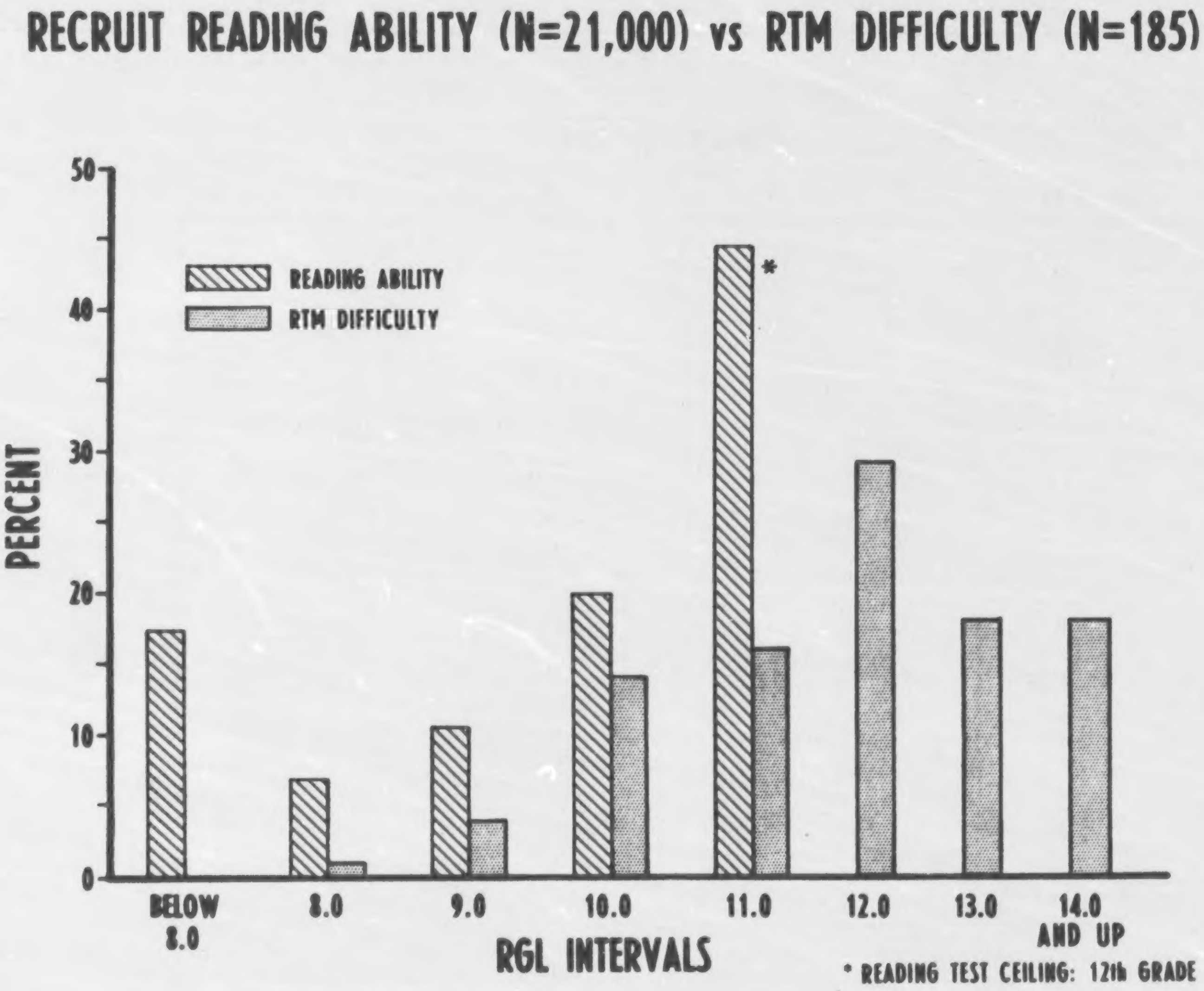


Figure 1. Comparison of recruit reading abilities and training manual difficulty as reported by Biersner.

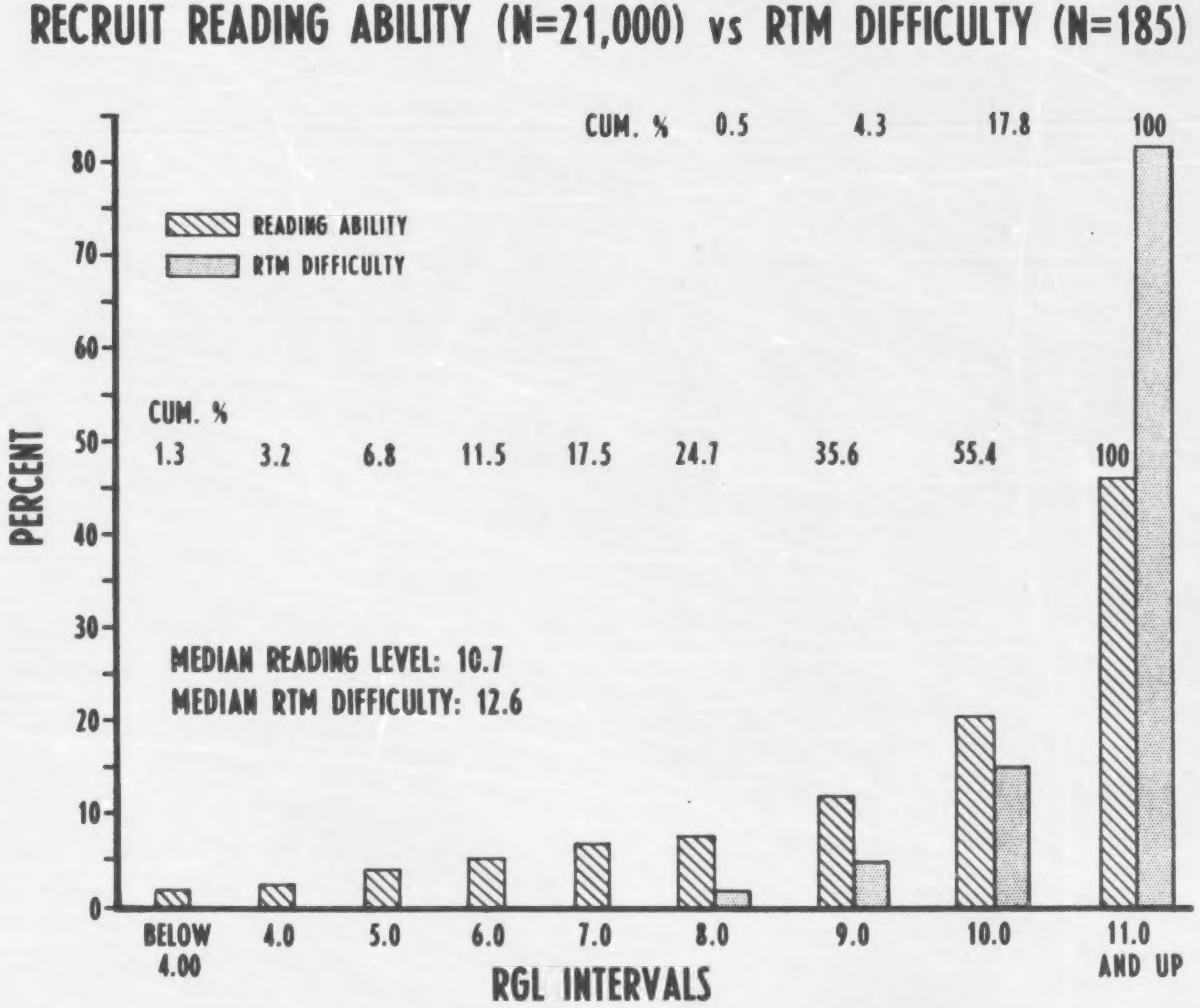


Figure 2. Comparison of recruit reading abilities and training manual difficulty of recent Navy recruits.

Technical Terms and Measurement of Difficulty

A major problem with most readability formulas (as alluded to above) is that they were developed and validated using either school children or civilian adults. There is good reason to believe that the abilities of a child reading at the 6th grade level (for example) and of a Navy man also reading at the 6th grade level are quite different. Again, the military environment itself may be expected to account for perhaps great differences in the words with which the person is familiar. This problem is even more acute when one is concerned with technical writing. Since word difficulty is usually indexed in terms of the length of the word, the technical terms encountered on the job will tend to inflate readability formula scores. Yet it is probably nonsense to assume that an electronic technician would be unfamiliar with the word "oscilloscope." Traditional formulas have not addressed this problem directly, and yet they must. It is analagous to the ongoing theme of this paper. The fact that a readability formula indexes a piece of writing as "easy" or "difficult" does not ensure that it will appear that way to the intended user. Two questions seem to be at issue here: to what degree do technical terms inflate the analysis (since they seldom exceed about 15 percent of the total number of words), and, if indeed they are a problem, by what means can the difficulty of technical writing be realistically indexed? The first of these questions is a basic empirical one, and should be answered before proceeding with R&D involving the second question. Additional formulas, and measurement considerations not involving formulas, will be discussed below, taking these R&D priorities into consideration.

The FORCAST and RIDE Formulas

Caylor, Sticht, Fox, and Ford (1973) made an attempt to allay the problem of using formulas derived from civilian samples in determining the difficulty of military writing. They examined a total of 15 structural variables in Army training literature, including words per sentence, number of independent clauses, number of one-syllable words, and total number of syllables. They were also dissatisfied with the traditional readability formulas for use with military writing. They state:

> The fact that the formulas have validity coefficients of about .70 for predicting the performance of school children on reading comprehension tests indicates that they account for roughly 50% of the variability in reading performance of children. It is likely that they may account for less variability in adult performance, especially since material containing large numbers of technical terms would increase the estimate of difficulty made by the readability formulas. (p. 6)

They therefore examined job-relevant materials from seven military occupational specialities (MOSs) in the Army, using scaled RGLs of this material (based on the cloze test) as a criterion, and computed regression equations with the 15 variables. They found that the number of one-syllable words alone correlated .86 with the criterion cloze scores, and that the addition of either one or two additional variables to the equation produced no significant increase in this correlation. They therefore included only this variable--number of one-syllable words--in their formula, which was termed "FORCAST." This formula became:

$$\text{FORCAST RGL} = 20 - \frac{\text{number of one-syllable words}}{10}$$

Based on their results, and considering that this vehicle is probably the most simple count to achieve reliably, the rationale for recommending the one-variable formula seems sound. It should be noted, however, that others have found that the addition of a sentence-length variable adds considerably to predictive power. Kincaid et al. (1975), for example, found that the addition of a sentence difficulty factor to the word difficulty variable increased the "coefficient of determination" (indicating the degree of shared variance between the predicted grade level and comprehension of the test passage) from 41.6 percent to 57.2 percent. This is a sizable increase in power, indicating that the findings of Caylor et al. (1973) might be reexamined to determine if the difference in power outweighs the ease of computation.

Carver (1973, 1974) also developed a single-variable formula, which he termed the RIDE scale (an acronym for Reading Input Difficulty Estimate scale). The variable used in this formula is simply the average number of letters per word (lpw) in a passage. There are five levels which index RIDE difficulty: Level 1 (up to 4.0 lpw), Level 2 (4.1 to 4.5 lpw), Level 3 (4.6 to 5.0 lpw), Level 4 (5.1 to 5.5 lpw), and Level 5 (5.6 lpw and above). This formula would appear to suffer from the same deficiency as the FORCAST formula. That is, it considers only word difficulty and ignores sentence length. Further, it remains to be seen how this formula would hold up with technical writing containing long but familiar terms, in view of the fact that it was validated on the Miller-Coleman passage (Aquino, 1969) using school children. This question should be tested to determine its validity.

Carver based his work on the findings of Bormuth (1966, 1969), who showed that the number of letters per word was one of the highest single correlates of the cloze difficulty of 330 100-word test passages. According to Bormuth (1969) only the Dale "Long List" correlated more highly as a single variable than number of letters per word (and that by only .006). It should be noted, however, that both these variables deal only with word length/familiarity, and do not touch upon sentence difficulty. It has been shown consistently that word length or familiarity accounts for a greater amount of variance than does sentence length. If one's major goal is to devise a prediction formula containing only one variable, then the rationale for the RIDE scale would be sound. It seems worthwhile, however, to pursue further the potential increase in predictive power with the addition of a sentence difficulty factor.

Bormuth (1969) did go beyond the examination of single variables in isolation. He conducted an extensive correlational study using approximately 170 different variables, of which lpw was one. As a first step, he determined the correlation of each of these variables singly with the scaled difficulty level of passages as referred to above. However, he then went on to determine, using multiple regression techniques, which of these variables in combination yielded the best predictive power. He arrived at four basic passage length formulas (each of which was computed against four different criteria). Two of these formulas (termed "unrestricted") employ large numbers of variables and are clearly unsuitable for other than very sophisticated automated computation. (Interestingly, neither of these basic formulas includes lpw as a variable.) To exemplify these formulas, and considering the fact that automation seems obvious to become the rule for the ultimate analysis of difficulty, the "short form" of the unrestricted formula will be described briefly. This formula consists of eight different variables plus the square of the two of these, totaling 10 entries in the formula. The eight basic variables are (1) number of Dale Long List words, (2) letters per minimal punctuation unit (and its square), (3) number of "referential repetition anaphora" (and its square), (4) number of numerical nouns, (5) number of derived adjectives, (6) number of common nouns, (7) number of relative clauses, and (8) number of "class inclusive anaphora." It seems obvious that the present state-of-the-art of readability analysis does not permit efficient application of such a complex formula. Its correlation with the mean cloze criterion (.874), however, is exceeded only by that of the even more complex basic unrestricted formula (.889). The question remains (even disregarding computer state-of-the-art) whether these correlations represent diminishing returns over the somewhat lower correlations of the much more simple manual and machine computation formulas described below.

Looking at the letters per word variable alone and comparing it to the two Bormuth formulas which include it and other variables, the results are as follows:

(1) Letters per word alone correlated -.721 with mean cloze scores. (For the purpose of this comparison only mean cloze score will be examined.)

(2) Letters per word, when combined with letters per minimal punctuation unit in the manual computation formula, correlated .808 with mean cloze score (and .833 in cross-validation[3]).

(3) Letters per word, when combined with number of Dale Long List words and words per sentence in Bormuth's machine computation formula, correlated .833 with the criterion (and .920 in cross-validation).

---

[3] Cross-validation was accomplished by applying the formulas to 20 passages of 275-300 words each taken from earlier work by Bormuth (1966).

Disregarding the fact that each of the correlations increased in cross-validation, the addition of new variables to the single lpw variable resulted in an increase in observed (or shared) variance of from 14-17 percent. It remains to be seen whether the additional difficulty and possible added unreliability in counting letters per minimal punctuation unit or words on the Dale Long List would offset this rather sizable increase in predictive power. Carver reports correlations of .93 and -.94 between the RIDE scale and the Dale-Chall and Flesch formulas, respectively, using the Miller-Coleman passages as criteria. He concludes that, for these passages, "there seems to be little difference between the predictive validity of Dale-Chall, Flesch, and RIDE" (Carver, 1974, p. 12).[4] These discrepant findings--both of the other formulas consider sentence length and word difficulty or familiarity--remain to be explained.

The Fog Count

The Fog Count, developed originally by McElroy (1953) for the Air Force, is perhaps the most simple two-variable formula to compute manually. It involves simply a count of the number of words of three or more syllables, the total number of words, and the number of sentences and entering them into a simple formula. Rather than discuss the original Fog Count it would seem more appropriate to address the "recalculated" Fog Count presented by Kincaid et al. (1975). According to these authors, the revised version is "very similar" to the original except that a different subtracted constant is used to redress a problem of over-estimation encountered with the original. The Kincaid et al. (1975) recalculation was accomplished using 569 Navy enlisted personnel reading material from Navy rate training manuals, at least some of which was technical in nature. In both the original and the revision, long familiar terms are treated specially to avoid the problem of their contributing erroneously to the difficulty of the material. For example, "General Eisenhower" (consisting of a total of seven syllables) is considered as one "easy" word for the purpose of calculating the Fog Count. Other units, such as numbers and some abbreviations, are treated in a similar manner. It would seem that a similar procedure could be applied in dealing with technical terminology known to be familiar to the user of a given piece of material. Here a problem enters that is related to the first of the basic questions posed above. Before the Fog Count (or any other formula) can be made to take account of such terms, the terms themselves must first be identified. This is not so easy as one might first suppose. In the field of electronics, for example, a basic dictionary of terminology for comprehension of written material could be obtained from an examination of books in the area. But to whom will this dictionary apply? Will it be sufficient to have only one such dictionary? If there is only one, will it apply equally as well to the difficulty of material designed for basic electronics training (e.g., "A" school) and the senior petty officer on the job? Considering these questions, the research involved would seem to be a difficult, but not insurmountable effort. And once the dictionary(ies) is (are) constructed, there is the question of whether the additional effort required in applying them is accompanied by a significant and practical gain in our ability to gear materials to the desired level of difficulty. With regard to the Fog Count, and presuming that a dictionary of technical terms proves worthwhile, the present advantage of easy manual calculation will probably reduced. The effective use of a dictionary of technical terms will virtually demand the automation of the predictive device. The Fog Count in manual form, however, might even then be of value for a "rough" difficulty analysis where automation is not available. Reference to a "short list" of, say the 25 most common terms, could be accomplished during a manual count, and such words treated as "easy" for the purpose of computation. Further reference to this problem will be made with regard to the Dale-Chall formula below and in the discussion of "production" of readable writing.

---

[4] Carver points out, however, on the basis of later work, that the Flesch and Dale-Chall formulas seem to be "consistently" better than others.

The Dale-Chall Formula

The Dale-Chall formula (Dale & Chall, 1948) is a two-variable formula which is based on average sentence length and the number of words in the passage not on the "Dale List of 3,000 Easy Words," (This "Dale List" was compiled by asking fourth graders which of a number of words they "knew.") To the best of this author's knowledge, the Dale-Chall (D-C) formula has not been validated in a military setting. It is discussed here because it has consistently been found to be highly valid when compared with other formulas in non-military settings, and because it offers a "model" for construction of a technical dictionary or lookup list such as discussed above. Again, presuming that the appropriate technical terms have been identified, such terms could possibly be simply added to the Dale List. Then, when encountered in text (since they are on "the list"), they would be considered "known" or "easy" and not contribute to the difficulty index of the passage. If research indicates that a technical dictionary does improve our prediction ability for military writing, the D-C would be recommended for use due to its "track record" of validity.

Other Methods of Measuring Difficulty

The above discussion has considered only the prediction of readability by basically "statistical" methods. Mention should also be made of two other methods for determining the readability and comprehensibility of written material. The first of these involves a judgment of the quality of material by persons who read a passage and compare it to some standard or to their own experience of job reading requirements. Carver (1974) examined this procedure using his "Rauding Scale." A major advantage of this method over that of formulas is that style, sentence structure, etc. can be assessed by the reader, a task which formulas are unable to carry out. In Carver's approach, judges who qualify on the "Rauding Scale Qualification Test" are asked to judge sample passages and categorize them according to a set of six "anchor passages." Using the average rating of three such judges, a grade level for the target passage is determined. In short, Carver found the correlations between Rauding Scale judgment and the RIDE, Flesch, and D-C formulas to range from .74 to .84. These moderately high correlations indicate that human judges are taking into account characteristics of writing similar to those used in calculating the formula scores. It seems equally obvious, however, that human judges are considering "something else" in their ratings; what this might be one can only surmise. Carver (1974) hypothesizes that the Rauding Scale reflects "the difficulty of the ideas or concepts in a passage," and indicates when "choppy sentences and inappropriately inserted little words" make material harder to read and understand (than a statistical estimate would indicate). The major disadvantage of the use of human judges in the assessment of readability and comprehensibility is simply availability--availability of time and of personnel. For such a procedure to be used effectively with the vast amount of material with which the Navy is concerned would seem to be prohibitive in the extreme.

Another method for determining the quality of written material is the "cloze test" referred to above in the context of readability criteria. There seems to be considerable disagreement as to what the cloze test actually measures. Some say it measures readability; others use the terms comprehensibility, comprehension, or understanding. Kincaid et al. (1975), for example, state that "subjects were tested for their understanding of the selected passages using the cloze procedure" (p. 3, emphasis added). Sticht (1975) says "the cloze test provides a valid measure of reading comprehension" (p. 20, emphasis added). A number of writers (including Taylor, 1953, who developed the procedure) refer to the cloze technique as a measure of readability.

If the operational definition of readability is taken to be the index provided by a readability formula, the cloze test will not consistently measure readability. The results of a cloze test on the "scrambled" version of "Mary Had A Little Lamb" (for example) would probably bear little resemblance to the low difficulty predicted by a formula. Clearly, this is an extreme case. When writing of fairly high quality is examined there is consistently a moderate to high correlation between cloze performance and readability as predicted by formula. But one would expect, as the general "quality" of writing decreases, this correlation would correspondingly drop. It does not seem wise to use as a measuring device such a "rubber yardstick."

A correlation is also consistently found between cloze scores and "comprehension" or "understanding." It is fairly well established that a cloze score of about 40-45 percent corresponds to a comprehension test score (after reading the material) of about 75 percent. But again, it can be shown that there are limiting factors. This author has conducted a relatively modest pilot study in which cloze tests were administered on a single passage of about 200 words. Subjects were then asked to state as precisely as possible what the passage was all about. Only 3 subjects out of the approximately 50 tested was able to provide even a "reasonable" approximation of the passage content. Yet the mean cloze score over all subjects was approximately 40 percent. This would seem to be at least an indication that the subjects were not "comprehending" what they were reading. It is anticipated that a follow-up study will examine the differences in performance when (1) a title descriptive of the content of the passage is provided, and (2) a traditional test of comprehension is administered following the cloze test. Again, since there are clearly limits beyond which the cloze-comprehension correlation will not hold up, should we not be more precise in speaking about cloze in general?

It is held here that cloze measures comprehensibility. Klare, Sinaiko, and Stolurow (1972) take a similar position when they say "cloze measures the relative comprehensibility or intelligibility of written material." To reiterate an earlier statement, readability and comprehensibility are not one and the same. Comprehensibility entails such factors as writing style, sentence structure, and expectancy for familiar words. The first two of these factors cannot be detected at all by available readability formulas. The latter can be to some extent, although the identification of those words which are familiar to a given audience has not yet been satisfactorily accomplished. Yet these factors (and undoubtedly others, as well) are critical to performance on a cloze test. Multiple "embedding," for example, would most likely cause cloze performance to deteriorate greatly. Consider the following sentence:

> "This is the malt the rat the cat the dog teased killed ate that lay in the house that Jack built."

In this example--a syntactically correct sentence with a relatively low level of difficulty according to formula--a person encountering it for the first time would likely have a very difficult time with a cloze version. Style--that undefinable characteristic of writing--will cause even greater problems. There seems to be little doubt that the smooth flow of ideas, with few departures from the "standard syntax", enhances the probability that the word which occurs to a reader for a given "set" will be the correct one. In short, it is considered that the cloze test should be considered a supplement to a readability formula. Given that a passage is readable (as predicted by a formula), the cloze performance of a set of subjects similar to the intended audience of the passage should ensure that the material is also comprehensible.

Assessing Job Reading Requirements

Formulas, judgments, and cloze scores all clearly play a role in determining the level of ability required by a user of written materials. They do not, however, give an indication of the reading ability required for a specific job or even of the reading required for that job. The Navy Personnel Research and Development Center (NAVPERSRANDCEN) and the Human Resources Research Organization (HumRRO) recently set out to seek answers to these important questions. This research effort involves the direct assessment of job incumbents' ability to read and understand job reading materials. It involves a structured interview of the job site in which the men are asked to report what specific reading they have undertaken in the previous day or so that was directly connected with their work. When a man reports that he used a specific portion of a specific publication to perform a specific job, he is asked to retrieve that same information in the presence of the interviewer. From this interview data, a Job Reading Task Test (JRTT) battery will be constructed, containing specific tasks requiring reading which it is known that job incumbents actually do. These tasks will be broken down into categories such as following written directions, use of reference materials, identifying main ideas, etc. The actual materials in each of these categories used on the job will be reproduced or used intact in the JRTT battery. The battery will then be normed on a sample of recruit personnel of known reading ability but not possessing job-specific skills and knowledges. Once the reading ability required for each of the tasks is determined, they will be submitted to job incumbents for judgments of the criticality of each and the frequency with which each is performed. The final product will be an inventory of various types of tasks and the reading level associated with each. The inventory can then be used for establishing general reading requirements for any Navy training or job situation. In short, the inventory will make it possible to determine the nature of reading tasks in any rating, the criticality of each, the frequency with which it is performed, and an estimate of the reading ability required in order to perform it.

It can be seen that in one sense, this is an alternative to the use of readability formulas in assessing the difficulty of materials designed for a specific job. It has the advantage, however, of determining reading abilities required for specific materials known to be used on the job, whereas using formulas, some (or perhaps much) of the material sampled is not actually used by job incumbents (such as "theory of operation" for example). Put another way, the fact that written material is "designed" for a particular job does not necessarily mean that it is required to perform that job. Needless to say, the JRTT method does not supplant the need for formulas, judgments, or cloze tests. When any new material is sent into the field, it is obviously necessary to ensure, as far as possible, that the material will be satisfactory in terms of readability and comprehensibility. Whether it will be used, these tests will not say. The ultimate tests of its usability must come from the man on the job.

In addition to providing information as to the reading ability required to perform particular jobs, it is hoped that information regarding "readership" will also emerge. That is, does a gap between reading ability and difficulty of required job reading materials influence the extent to which men report using those materials. In discussing a similar study for the Army, Sticht (1975) reports that "the more able the reader, the greater the reported use of printed materials" (p. 52).

Summary

Summarizing the area of prediction of readability and comprehensibility, it seems clear that not just one, but a number of approaches should be used to promote high quality written materials. Readability is a basic characteristic and, in its operational sense, must be assessed according to a formula. But this is not sufficient. The comprensibility of the material must be assured by means of cloze tests or human judgments, if possible. Which formula should be used to predict readability, and which technique(s) should be used to assess comprehensibility are still matters open to question. And finally, to ensure that job reading materials are matched to reading abilities of job personnel, a technique such as the JRTT Inventory being developed by NAVPERSRANDCEN and HumRRO should be considered for Navy-wide use.

Production of Readable Writing

General

Predicting the difficulty of written material "after the fact" is a major problem, but equally or more important is the problem of producing "readable writing" in the first place. For this latter task there is no "formula" which one can directly apply. Writing is an ART as well as a SKILL. While it is probably impossible to tell someone how to "do art" there should be some techniques by which the writer's skill can be enhanced. This is the concern of this section, and in general, can be conceptualized as a three-part problem:

(1) Can a writing style guide be developed which will serve as a "job performance aid" for good writing?

(2) Given that a writer is both skilled and artistic in his work, can Military Specifications be provided him so that his product conforms to our needs?

(3) How can we best provide the writer with "tools" to ensure that specified standards are being met?

Readability formulas play a part in the production process, but as pointed out so clearly by Gunning (1968), "formulas are tools, not rules . . . warning systems, not formulas for writing." Production is not simply the opposite side of the coin from prediction, although the two processes are inextricably related. The two processes can be thought of as iterative: one writes, then assesses the difficulty of the writing, then rewrites, then reassesses, etc. The interaction between the two will be examined more thoroughly below in the contexts of specific writing aids.

Research in the area of readable writing does give some cause for optimism, but for caution as well. There have been a fairly large number of studies which have shown clearly that experimental manipulation of writing variables can make a significant difference in comprebensibility. In general, these experimenters (many of them conducting Ph.D dissertation research) have taken original passages and written easier and/or harder versions as indexed by readability formula scores. In examining these sources, the problem is one of specifying precisely what was changed to make the readability different. Most have reported that they modified word and sentence difficulty and little else, presumably because these are the two variables common to most readability formulas. Most writers would argue, however, that such changes are either not sufficient or are too simple-minded for such a complex task. It is here that a word of caution is in order. There seems to be little doubt that mechanical shortening of words and sentences will do little, if anything, other than improve readability formula scores, with scant chance of improving the reader's comprehension. At the extreme, it is clear that a writer with ulterior motives could certainly produce writing that gets a better readability score and yet be less comprehensible than the original version. It seems clear, then, that in the research studies mentioned above, the modification of surface variables (i.e., word length, sentence length, etc.) have incidentally resulted also in changes to deeper (or causal) variables. This is a research question which is virtually untouched and yet which is critical to the understanding of how to make writing more comprehensible.

Style Guides

The first facet of the problem of readable writing as mentioned above--style guides to aid the writer--would profit immensely from the research just suggested. The guidelines put forth by existing style manuals are based largely on intuition with little, if any, empirical base. It may be that such manuals are of no help to writers. They have existed in profusion in our libraries for years, and yet we are still suffering from inferior written materials. One must conclude from this that such guides are either not used by writers of military materials or that the information they contain is inadequate to the task of telling "how to write." Information provided this author by Dr. G. R. Klare (personal communication) indicates that the problem lies, at least in part, with the latter of these possibilities. Klare reports that in reviewing 15 source books (10 written specifically for technical writers) the agreement among authors as to specific suggestions was quite low. The suggestion "use short words" (for example), which one might expect all authors to agree upon, was mentioned in only two of the 15 books sampled. Outright disagreement was found for such alternatives as "be concise" versus "be complete," and "keep paragraphs short" versus "vary paragraph length." In view of this apparent uncertainty among "experts" on writing, it is not surprising that much of our material is unacceptably difficult. It is felt that progress is being made in this area with the publication of style guides specifically for the military writer. Under contract with the Naval Sea Systems Command, for example, BioTechnology, Incorporated produced a guide titled "Requirements and Criteria for Improving Reading Comprehension of Technical Manuals" (Post & Price, 1974). This guide contains 17 "tests" for improving the quality of technical writing in the three areas of Organization, Technical Communication, and Readability. The seven tests dealing specifically with readability and comprehensibility, in abbreviated form, are as follows:[5]

---

[5] This manual assumes that the writer is aiming at the ninth grade level; this should be considered in interpreting the tests. An answer of "yes" to any question indicates adequate readability (at ninth grade) on that test.

1. Heading Review. Do approximately one-half of the subparagraphs have headings? Is material within paragraphs consistent with its heading?

2. Topic Sentence Check. Is the heading clear? Does the heading cover about three or four topic sentences or key points?

3. Words Per Paragraph Count. Do paragraphs average no more than 45 to 60 words? Are key points highlighted if the paragraph must be longer?

4. Words Per Sentence Count. Do the sentences average 20 words or fewer? Have compound sentences and complex sentences been avoided?

5. Syllables Per Word Count. Does the material average about 1 1/2 syllables per word? Have short words been used whenever possible?[6]

6. Equipment Nomenclature Count. Is any unfamiliar nomenclature either defined in the text or called out on an accompanying pictorial?

7. Layout Review. Has double-column format been used? Is each graphic contiguous with the text in which it is discussed or referenced?

It is believed that this manual (which also contains specific guidelines for correcting deficiencies, a large number of "rewrite practices," and generous illustrations of "good" and "bad" material) can be of valuable assistance to the technical writer. However, it is vital that it be experimentally tested to verify this assumption. Work is currently underway which involves the use of this guide in rewriting a technical manual and comparing the rewritten version with the original in terms of its effectiveness.

Military Specifications

A second major area of concern in the production of readable writing is that of communicating the needs of the user to the technical writer. This is a question of the clarity, conciseness, and completeness of Military Specifications (Mil-Specs) and Military Standards (Mil-Stds). With regard to readability and comprehensibility, it is the opinion of this writer that present Mil-Specs and Mil-Stds offer little useful guidance for the writer. Other matters aside, it is clearly apparent that no test of readability or comprehensibility has been applied to *these* documents themselves. One sentence picked more or less at random from one basic Mil-Spec contained 47 words, of which 21 words were composed of three or more syllables. Other similar examples abound. But more specifically, with regard to specification of readability and comprehensibility standards, most or all such documents are woefully inadequate. One, for example, states that "As a general guide, the level of writing should be for a high school graduate having specialized training as a techician in military training courses."[7] Even those of us in the field of readability/comprehensibility (much less the engineer/writer) don't really know what a typical high school graduate "looks like." And taken literally, manuals written to this prescription (if it were possible) would automatically be excluded from use in at least some training courses, because the trainee would not yet have the "specialized training" required to understand the manual. It would be comparable to the employment advertisements which specify "only experienced need apply," when the "experience" can be gained only through the employment.

---

[6] The authors suggest that since manuals deal with technical terms which cannot be eliminated, these terms should not be included in the count of syllables. This procedure should, however, be experimentally tested.

Another, more basic Mil-Spec gives somewhat more explicit directions. It states: "Narrative text (those pages that consist of not less than 200 words in consecutive sentences per page) shall conform to the following readability standards: The average sentence length (ASL) shall not exceed 20 words . . . The average word length shall not exceed 1.60 syllables . . . The percent personal sentences (PPS) shall not be less than 15 percent of the total."[8] The particular figures specified are reasonable; if such standards were achieved, the readability formula score would be at approximately the 9th or 10th grade level. But, one might wonder if some writers read any further than where it says " . . . 200 words in consecutive sentences . . ." Certain persons responsible for producing technical manuals have been heard to say that "by definition" their writing never has more than 200 words per page in consecutive sentences. This is not to imply that producers or writers of manuals intentionally write poorly. Quite the contrary, one suspects that they take pride in their work just as any professional does. But the fact remains that much of our technical writing is too difficult for many of the intended users. More specific guidance, and perhaps more rigid quality control, would appear to be necessary conditions for rectifying this problem. And accompanying such specifications, it is felt that we must provide the writer with tools to meet our standards--the third general problem in the production of readable writing.

Tools for Readable Writing

What tools could we provide for writers? In general the answer involves the various methods of assessing readability and comprehensibility discussed in the previous section. As indicated earlier, prediction and production go hand in hand. Figure 3 illustrates the iterative process that is involved. Included in this figure are some examples of the tools that might be provided the writer in order to produce an acceptable product. Basically, we see the writer producing his "first draft," relying at the outset primarily on intuition in predicting its difficulty. At some point, however, a more precise assessment of the writing should take place. This might take the form of a readability formula (which the writer will be trained to use) computed manually. As pointed out earlier, the formulas which are currently available from RIDE and FORCAST to the probably more precise Flesch (Kincaid) and Dale-Chall, are all relatively simple to calculate by hand. As our understanding of the problems specific to military writing (e.g., long, but familiar technical terms) become better understood, and as cost-effective and more efficient methods of analyzing readability and comprehensibility are developed, it would seem appropriate to automate the process of prediction to the greatest possible extent. Several possibilities present themselves in this regard, ordered in successively more sophisticated techniques.

---

[7] MIL-M-24100B, Manuals, Technical: Functionally Oriented Maintenance Manuals (FOMM) for Equipment and Systems, January 1974.

[8] MIL-M-63000C (TM), Manuals, Technical: General Requirements for Manuscripts, December 1960.

One such technique is the Navy Automated Counter (NAC) developed for use in the CNETS analysis of Navy Rate Training Manuals (RTMs) (Biersner, 1975; Bunde, 1975). To reiterate briefly, this device is composed of a stylus which, when pressed to a working surface, trips a microswitch, which in turn activates a counter. Used for counts of syllables, words, and sentences for the Flesch and Kincaid formulas, it would be equally as useful for computing virtually any other formula. This could be a cost-effective aid which writers would probably find more acceptable than using simply paper, pencil, and fingers. This "acceptability" would probably give greater assurance that accurate counts are in fact being made in the field. Bunde (1975), in a test of this device, found that it reduced the time to compute the Flesch formula by 46 percent over manual counts, with equivalent reliability. From the point of view of this writer, such a device might be consdered as an interim technique. It is recommended, however, that more sophisticated procedures be planned for future use.

At a somewhat higher level of sophistication are those readability analysis techniques which utilize a specially modified electric typewriter. These are the Automated Readability Index (ARI) and the Reading Ease Assessment Device (READ). Given material to be analyzed during any typing stage of the production process, these devices (which are similar in design) allow the writer to determine the difficulty of his writing as it is typed. Essentially, these devices take account of the number of "strokes" of the machine to count the number of letters per word and the number of words per sentence to be entered into a modified Flesch RE formula. They both require typing skill, however, which may be a drawback for some writers. In addition, their cost may be prohibitive (although not exorbitant) for situating at the many locales at which manuals are produced. If however, the cost is warranted, and if skilled personnel are available for their use, they represent a step advanced from the manual counting device discussed above. But in addition to the above factors, these devices are limited in that they provide only readability scores. As will be seen below, improvement of technical writing overall actually demands much more.

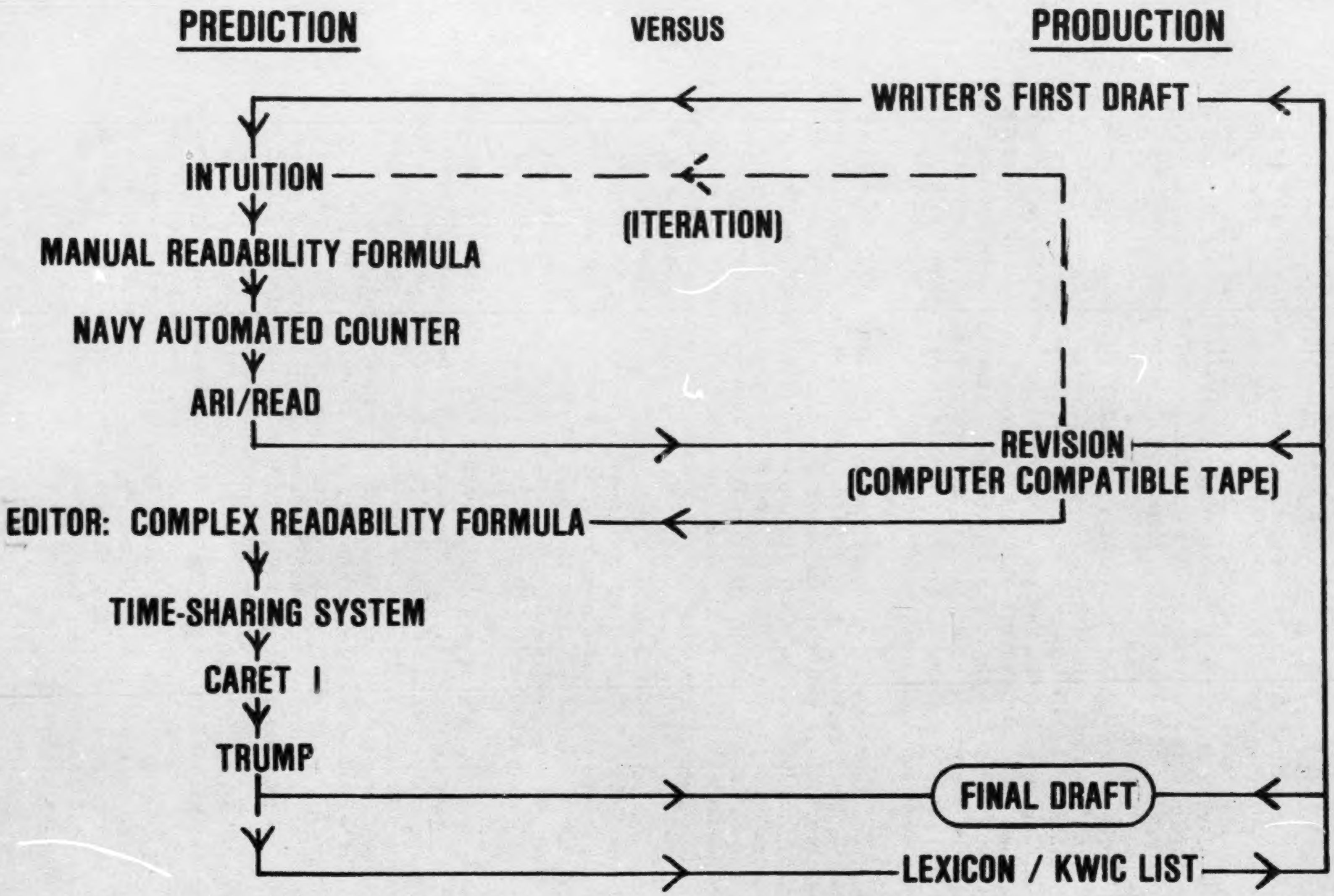

PREDICTION
VERSUS
PRODUCTION
WRITER'S FIRST DRAFT
INTUITION
(ITERATION)
MANUAL READABILITY FORMULA
NAVY AUTOMATED COUNTER
ARI/READ
REVISION
(COMPUTER COMPATIBLE TAPE)
EDITOR: COMPLEX READABILITY FORMULA
TIME-SHARING SYSTEM
CARET I
TRUMP
FINAL DRAFT
LEXICON / KWIC LIST

Figure 3. Prediction of difficulty versus production of readable writing.

At a higher level of potential, let us consider two possible systems--one marketed by Scientific Time-Sharing, Incorporated and the other developed at Harvard University.[9]

The first of these systems incorporates a modified RE computation, in which a sample of material can be input to a computer terminal which is connected via telephone line to a central computer. Material is typed at a remote terminal and, on signal, an almost instantaneous analysis of the writing is printed out. A seemingly advantageous characteristic of this technique is its ability to reduce the impact of long technical terms on the analysis. When inputting the material, if a familiar but long technical term is encountered, it is counted as any other word at its first appearance--both as a discrete word and as a polysyllabic word. When subsequently encountered, it can be enclosed in brackets which excludes that occurrence from the overall analysis. The effectiveness of this technique, however, depends entirely on the ability of the typist to recognize those words which are of such interest and bracket them in an efficient manner. Unlike the ARI or READ devices, however, it is considered that the time-sharing system offers the potential for such features as storing of technical terms in a lexicon and future automatic exclusion of these terms from the analysis. It is considered that future R&D effort might well include the investigation of such potential.

The second of these computerized techniques is referred to as a Computer Aided Revising, Editing, and Translating system (CARET I), and was developed by Klare et al. (1969). Like the time-sharing technique discussed above, CARET I also provides readability analyses after the material to be assessed has been input. It is felt that CARET I has at least two distinct immediate advantages over the former system. First, it provides not one, but five separate analyses, including the Flesch, the Fog Index, and the Farr-Jenkins-Paterson revision of the Flesch formula. Secondly, it provides a triple-spaced printout of the inputted material, indicating for each word the number of syllables, and for each sentence the number of words. The editor/writer can therefore adjust (as necessary) word length and/or sentence length, entering alternatives directly on the printout. And, like the time-sharing system, it is interfaced with a large computer which would permit the compilation of a lexicon and many other possibilities to assist not only in prediction of difficulty, but in writing and rewriting. Figure 3 shows how such a lexicon (as well as a Key-Word-in-Context file) might "feed back" into the various stages of writing. To exemplify the potential of such a system, the entire text of (for example) an electronics manual could be analyzed for the frequency of occurrence of key words at the same time that the material is input for readability analyses. Then not only those words in the lay language, but also technical terms, could be cataloged according to their frequency and this information fed back to writers for consideration either in rewriting or in original writing of new material in the same subject matter area. Optimally, it would appear that the best features of both the preceding systems could be combined to form a procedure that is efficient, effective, and cost-beneficial.

---

[9] No bias in favor of these specific developers is intended. Other similar (and perhaps better) systems may well exist and not have come to the attention of this author.

An even more sophisticated system would be one such as the Navy's Technical Review and Update of Manuals and Publications system (TRUMP). This system is currently undergoing test and evaluation at the Naval Air Rework Facilitty, Jacksonville, Florida. TRUMP is technologically more advanced than the others mentioned, basically because of its Optical Character Recognition (OCR) and automatic photocomposition capabilities. This makes it possible to electronically scan material to be input, resulting in a reported throughput rate of potentially hundreds of times greater than with keypunching. While no readability analysis for TRUMP has as yet been implemented, it would seem to have the potential for virtually any such analysis as well as all other capabilities mentioned in connection with previous systems. Of particular interest for Navy manuals, TRUMP can automatically process illustrations and complex tables as well as running text. Upon completion of the ongoing T&E of TRUMP, it should be considered as a central feature of any future R&D effort in the rea of readability and comprehensibility.

Summary

The underlying theme of this paper has been a plea for reformulation of the basic rationale underlying research in readability and comprehensibility. It recommends that future research be directed at determining which of the many variables involved in the readability and comprehensibility area stand in causal relationships with the ability of persons to comprehend the written word. And finally it suggests that present readability formulas, with perhaps some modification, are acceptable when only readability is being considered, but that a means must be found to ensure that materials are also satisfactorily comprehensible. It seems that the time has come to orient our efforts toward implementation of what is already known in the field, and to provide for the military writer the means to achieve an acceptable product considered in terms of the needs of the user.

References

Aquino, M. R. The validity of the Miller-Coleman readability scale. *Reading Research Quarterly*, Spring 1969, 342-357.

Biersner, R. J. *Reading grade levels of Navy rate training manuals and non-resident career courses* (Tech. Rep. 2-75). Pensacola, Florida: Chief of Naval Education and Training Support, May 1975.

Bormuth, J. R. Readability: A new approach. *Reading Research Quarterly*, 1966, *1*, 79-131.

Bormuth, J. R. *Development of readability analysis* (Final Report, Project No. 7-0052). Washington, D.C.: Bureau of Research, U. S. Office of Education, 1969.

Bunde, G. R. *An effectiveness evaluation between manual and automated readability counting techniques* (Tech. Rep. 5-75). Pensacola, Florida: Chief of Naval Education and Training Support, August 1975.

Carver, R. P. *New techniques for measuring and improving reading comprehension* (Tech. Rep. 1). Arlington, Virginia: Office of Naval Research, February 1973.

Carver, R. P. *Improving reading comprehension: Measuring readability* (Tech. Rep.). Arlington, Virginia: Office of Naval Research, May 1974.

Caylor, J. S., Sticht, T. G., Fox, L. C., & Ford, J. P. *Methodologies for determining reading requirements of Military Occupational Specialties* (Tech. Rep. 73-5). Arlington, Virginia: Human Resources Research Organization, March 1973.

Curran, T. E. *Review of technical manual readability and comprehensibility* (Tech. Rep.). San Diego, California: Navy Personnel Research and Development Center (in preparation).

Dale, E., & Chall, J. S. A formula for predicting readability. *Educational Research Bulletin*, 1948, *27*, 37-54.

Flesch, R. F. A new readability yardstick. *Journal of Applied Psychology*, 1948, *32*, 221-233.

Gunning, R. *The technique of clear writing* (Rev. Ed.). New York: McGraw-Hill, 1968.

Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. *Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy enlisted personnel*. (Tech. Rep. 8-75). Millington, TN: Naval Technical Training Command, February 1975.

Klare, G. R. *The measurement of readability*. Ames, Iowa: Iowa State University Press, 1963.

Klare, G. R., Rowe, P. P., St. John, M. G., & Stolurow, L. M. *A first version of a computer-aided revising, editing, and translating system (CARET I)* (Tech. Rep. B.2.1). Cambridge, MASS: Harvard Computer-Aided Instruction Laboratory, 1969.

Klare, G. R., Sinaiko, H. W., & Stolurow, L. M. Cloze procedure: A convenient readability test for training materials and translations. *International Review of Applied Psychology*, 1972, *21*, 77-106.

Klare, G. R. Assessing readability. Reading Research Quarterly, 1974-1975, 10,(1).

Klare, G. R. A manual for readable writing. Glen Burnie, Maryland: REM Company, 1975.

McElroy, J. The Fog Count. Library Science, 1953, 4, 26-30.

Post, T. J., & Price, H. E. Requirements and criteria for improving reading comprehension of technical manuals (Tech. Rep.) Washington, D. C.: Naval Sea Systems Command, November 1974.

Sticht, T. G. (Ed.). Reading for working: A functional literacy anthology. Alexandria, Virginia: Human Resources Research Organization, 1975.

Taylor, W. L. Cloze procedure: A new tool for measuring readability. Journalism Quarterly, 1953, 40, 415-433.

Reference Note

1. Duffy, T. M., & Nugent, W. Reading ability and attrition. San Diego: Navy Personnel Research and Development Center, Unpublished Manuscript, 1975.